

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 11:39:30 ; Search time 32 seconds
(without alignments)
33.313 Million cell updates/sec

Title: US-09-660-302C-1
Perfect score: 8
Sequence: 1 xxxxxxxx 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 0 | 0.0 | 1 | 20 AAY46652 | Immunogenic peptid |
| 2 | 0 | 0.0 | 1 | 22 ABB56870 | Human SNP related |
| 3 | 0 | 0.0 | 1 | 22 ABB66809 | Drosophila melanog |
| 4 | 0 | 0.0 | 1 | 22 ABB66810 | Drosophila melanog |
| 5 | 0 | 0.0 | 1 | 22 ABG02941 | Novel human diagno |
| 6 | 0 | 0.0 | 1 | 22 AAW97643 | Human peptide #918 |
| 7 | 0 | 0.0 | 1 | 22 AAW97834 | Human peptide #110 |
| 8 | 0 | 0.0 | 1 | 22 AAW97974 | Human peptide #124 |
| 9 | 0 | 0.0 | 1 | 22 AAW98354 | Human peptide #162 |
| 10 | 0 | 0.0 | 1 | 22 AAW98447 | Human peptide #172 |

| | | | | | |
|----|---|-----|---|-------------|--------------------|
| 11 | 0 | 0.0 | 1 | 22 AAM53218 | Human nonconservat |
| 12 | 0 | 0.0 | 1 | 22 AAM53219 | Human nonconservat |
| 13 | 0 | 0.0 | 1 | 22 AAM53290 | Human nonconservat |
| 14 | 0 | 0.0 | 1 | 22 AAM53291 | Human nonconservat |
| 15 | 0 | 0.0 | 1 | 22 AAM53328 | Human nonconservat |
| 16 | 0 | 0.0 | 1 | 22 AAM53329 | Human nonconservat |
| 17 | 0 | 0.0 | 1 | 22 AAM45230 | H11 binding site c |
| 18 | 0 | 0.0 | 1 | 22 AAG99966 | ERA binding domain |
| 19 | 0 | 0.0 | 1 | 22 AAG99983 | ERA binding domain |
| 20 | 0 | 0.0 | 1 | 22 AAG99987 | ERA binding domain |
| 21 | 0 | 0.0 | 1 | 22 AAG99988 | ERA binding domain |
| 22 | 0 | 0.0 | 1 | 22 AAM00011 | ERA binding domain |
| 23 | 0 | 0.0 | 1 | 22 AAM00011 | ERA binding domain |
| 24 | 0 | 0.0 | 1 | 22 AAM00013 | ERA binding domain |
| 25 | 0 | 0.0 | 1 | 22 AAM00016 | ERA binding domain |
| 26 | 0 | 0.0 | 1 | 22 AAG98026 | Human SNP associat |
| 27 | 0 | 0.0 | 1 | 22 AAG98134 | Human SNP associat |
| 28 | 0 | 0.0 | 1 | 22 AAB91029 | Thyrotropin releas |
| 29 | 0 | 0.0 | 1 | 22 AAB91546 | Endothelins and re |
| 30 | 0 | 0.0 | 1 | 22 AAB91665 | Oploid peptide SEQ |
| 31 | 0 | 0.0 | 1 | 22 AAB91739 | Oploid peptide SEQ |
| 32 | 0 | 0.0 | 1 | 22 AAB91892 | Apoptosis related |
| 33 | 0 | 0.0 | 1 | 22 AAB92150 | Polypeptide SEQ ID |
| 34 | 0 | 0.0 | 1 | 22 AAB92392 | Miscellaneous pept |
| 35 | 0 | 0.0 | 1 | 23 ABG63437 | Human albumin fusi |
| 36 | 0 | 0.0 | 1 | 23 ABG63439 | Human albumin fusi |
| 37 | 0 | 0.0 | 1 | 23 ABG63578 | Human albumin fusi |
| 38 | 0 | 0.0 | 1 | 23 ABG63753 | Human albumin fusi |
| 39 | 0 | 0.0 | 1 | 23 ABG63997 | Human albumin fusi |
| 40 | 0 | 0.0 | 1 | 23 ABG64298 | Human albumin fusi |
| 41 | 0 | 0.0 | 1 | 23 ABG64370 | Human albumin fusi |
| 42 | 0 | 0.0 | 1 | 23 ABG64797 | Human albumin fusi |
| 43 | 0 | 0.0 | 1 | 23 ABB78513 | GAGP hydroxyprolin |
| 44 | 0 | 0.0 | 1 | 23 ABB78514 | GAGP hydroxyprolin |
| 45 | 0 | 0.0 | 1 | 23 ABB78515 | GAGP hydroxyprolin |

ALIGNMENTS

RESULT 1
AAY46652
ID AAY46652 standard; Peptide: 1 AA.
XX
AC AAY46652;
XX 01-DEC-1999 (first entry)
XX Immunogenic peptide having a human leukocyte antigen binding motif #1263.
DE Human leukocyte antigen; binding: immunogenic; glycoprotein; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX Synthetic.
OS Homo sapiens.
XX WO9945954-A1.
PN 16-SEP-1999.
PD
XX 13-MAR-1998; 98WO-US05039.
PF 13-MAR-1998; 98WO-US05039.
PR 13-MAR-1998; 98WO-US05039.
PA (EPIM-) EPIMMUNE INC.
XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
PI WPI: 1999-551214/46.
XX

PT New immunogenic peptides with HLA binding motif, useful in treatment
 XX and diagnosis of cancers and viral diseases
 PS Claim 1; Page 80; 150pp; English.
 XX
 CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) Class I (also
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
 CC response against the antigen from which the peptide is derived.
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
 CC normally induced by an antigen in the form of a peptide fragment bound
 CC to a HLA molecule, rather than the intact foreign antigen itself, and
 CC are particularly important in tumour rejection and in fighting viral
 CC infections. The peptides are therefore useful therapeutically to treat
 CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also
 CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polynucleotides encoding the immunogenic peptides are also useful
 CC therapeutically and for immunisation as above.

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 20; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 Y 1

RESULT 2
 ABB56870
 ID ABB56870 standard; Peptide; 1 AA.
 XX
 AC ABB56870;
 XX
 DT 05-MAR-2002 (first entry)
 XX
 DE Human SNP related amino acid sequence SEQ ID NO:1435.
 XX
 KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
 KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
 KW autoimmune disease; inflammation; cancer; nervous system disease;
 XX infection; polymorphic protein.

OS Homo sapiens.
 XX
 PN WO200138586-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32311.
 XX
 PR 24-NOV-1999; 99US-0167383.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI; 2001-355949/37.

XX Isolated human nucleic acids comprising one or more single nucleotide
 XX polymorphisms, useful for treating a subject suffering from a
 PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
 PT sequence polymorphism -

XX

PS Claim 1; Page 666; 674pp; English.

XX

CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
 CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
 CC to ABB56903 represent human peptides encoded by some of the SNP
 CC oligonucleotides. The sequences from the present invention can have
 CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
 CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
 CC and antibodies from the present invention can be used for treating a
 CC subject suffering from, at risk for, or suspected of, suffering from a
 CC pathology ascribed to the presence of a sequence polymorphism. The
 CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
 CC the nervous system, and infection by pathogenic microorganisms. The SNPs
 CC are also useful for determining which forms of a characterised
 CC polymorphism are present in individuals. The antibodies may be used in
 CC the detection, quantitation and/or cellular or tissue localisation of a
 CC polymorphic protein (e.g., for use in measuring levels of the
 CC polymorphic protein within appropriate physiological samples).

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 S 1

RESULT 3
 ABB66809
 ID ABB66809 standard; Protein; 1 AA.
 XX
 AC ABB66809;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 27219.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL10912.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -

PS Disclosure; SEQ ID NO 27219; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ

Sequence 1 AA:

Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 H 1

RESULT 4
ABB66810
ID ABB66810 standard; Protein; 1 AA.

XX ABB66810;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 27222.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL10913.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure; SEQ ID NO 27222; 2lpp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1 AA:

Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Qy 1 X 1

Db 1 H 1

RESULT 5
ABG02941
ID ABG02941 standard; Protein; 1 AA.

XX ABG02941;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #2932.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS67128.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

XX Claim 20; SEQ ID NO 33300; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1 AA:

Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 M 1

RESULT 6
AAM97643
ID AAM97643 standard; Peptide: 1 AA.
AC AAM97643;
XX
XX 24-JAN-2002 (first entry)
DT
DE
DE Human peptide #918 encoded by a SNP oligonucleotide.
XX
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
XX Homo sapiens.
XX
XX WO200147944-A2.
PN
XX
XX 05-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-US35498.
XX
XX 28-DEC-1999; 99US-0173419.
PR
XX 27-DEC-2000; 2000US-0173419.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shimkets RA, Leach M;
XX
XX WPI; 2001-465210/50.
XX
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
PT
XX Disclosure; Page 3869; 4143pp; English.
PS
XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 X 1
Db 1 C 1

RESULT 7
AAM97834
ID AAM97834 standard; Peptide: 1 AA.
XX
XX AAM97834;
AC
XX 24-JAN-2002 (first entry)
DT
XX Human peptide #1109 encoded by a SNP oligonucleotide.
DE
XX
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
XX Homo sapiens.
XX
XX WO200147944-A2.
PN
XX
XX 05-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-US35498.
XX
XX 28-DEC-1999; 99US-0173419.
PR
XX 27-DEC-2000; 2000US-0173419.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shimkets RA, Leach M;
XX
XX WPI; 2001-465210/50.
XX
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
PT
XX Disclosure; Page 3911; 4143pp; English.
PS
XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 X 1
Db 1 E 1

RESULT 8
AAM97974
ID AAM97974 standard; Peptide: 1 AA.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX Homo sapiens.
OS
XX
XX WO200147944-A2.
PN
XX
XX 05-JUL-2001.
PD
XX
XX 28-DEC-2000; 2000WO-US35498.
PF
XX
XX 28-DEC-1999; 99US-0173419.
PR
XX 27-DEC-2000; 2000US-0173419.
PR
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shimkets RA, Leach M;
XX WPI; 2001-465210/50.
XX
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
CC oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
PT
XX Disclosure; Page 4045; 4143pp; English.
PS
XX
XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX

Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 L 1

RESULT 11

AAM53218
ID AAM53218 standard; Peptide; 1 AA.

XX

AC AAM53218;

XX

DT 09-NOV-2001 (first entry)

XX

DE Human nonconservative amino acid changing SNP related peptide SEQ:6913.

XX

KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;

KW protein therapy; vaccine; probe; diagnostic assay; detection;

KW quantitation; restorative therapy; polymorphic.

XX

OS Homo sapiens.

XX

PN WO200140521-A2.

XX

PD 07-JUN-2001.

XX

PF 30-NOV-2000; 2000WO-US32758.

XX

PR 30-NOV-1999; 99US-0168138.

PR 29-NOV-2000; 2000US-0726173.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Leach M;

XX

DR WPI; 2001-356160/37.

XX

XX Polymorphic nucleic acid sequences, useful in genetic testing and

PT therapy -

XX

PS Claim 29; Page 2619; 2653pp; English.

XX

XX AAT73060 to AAT79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.

XX Sequence 1 AA;

SQ

Query Match 0.0%; Score 0; DB 22; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 H 1

RESULT 12

AAM53219

ID AAM53219 standard; Peptide; 1 AA.

XX

AC AAM53219;

XX

DT 09-NOV-2001 (first entry)

XX

DE Human nonconservative amino acid changing SNP related peptide SEQ:6914.

XX

KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;

KW protein therapy; vaccine; probe; diagnostic assay; detection;

KW quantitation; restorative therapy; polymorphic.

XX

OS Homo sapiens.

XX

PN WO200140521-A2.

XX

PD 07-JUN-2001.
 XX 30-NOV-2000; 2000WO-US32758.
 PF 30-NOV-1999; 99US-0168138.
 PR 29-NOV-2000; 2000US-0726173.
 XX (CURA-) CURAGEN CORP.
 XX Shimkets RA, Leach M;
 XX WPI: 2001-356160/37.
 XX Polymorphic nucleic acid sequences, useful in genetic testing and
 therapy -
 Claim 29; Page 2619; 2653pp; English.
 XX AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
 sequences (I), which contain single nucleotide polymorphisms (SNPs).
 CC AAM53114 to AAM53329 represent peptides related to human polymorphic
 CC polynucleotide sequences. The sequences can be used in gene and protein
 CC therapy, and in vaccine production. (I) and the polypeptides encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate expression of polymorphic polypeptides.
 CC For example, (I) may be used to treat disorders by rectifying mutations
 CC or deletions in a patient's genome that affect the activity of
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patients own production of polypeptide. Additionally, (I) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids
 CC in samples, and therefore which patients may be in need of restorative
 CC therapy. The polypeptides encoded by (I) may be used as antigens in the
 CC production of antibodies specific for polymorphic polypeptides. The
 CC antibodies may also be used to down regulate expression and activity.
 CC The antibodies may also be used as diagnostic agents for detecting the
 CC presence of polymorphic polypeptides in samples.
 XX Sequence 1 AA;
 SQ Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 H 1
 RESULT 13
 AAM53290
 ID AAM53290 standard; Peptide; 1 AA.
 XX AAM53290;
 AC 09-NOV-2001 (first entry)
 XX Human nonconservative amino acid changing SNP related peptide SEQ:6985.
 DE Human; single nucleotide polymorphism; SNP; genome; gene therapy;
 KW protein therapy; vaccine; probe; diagnostic assay; detection;
 KW quantitation; restorative therapy; polymorphic.
 XX Homo sapiens.
 OS WO200140521-A2.
 PN 07-JUN-2001.
 XX 30-NOV-2000; 2000WO-US32758.
 PF 30-NOV-1999; 99US-0168138.
 PR 29-NOV-2000; 2000US-0726173.
 XX (CURA-) CURAGEN CORP.
 XX Shimkets RA, Leach M;
 XX WPI: 2001-356160/37.

XX (CURA-) CURAGEN CORP.
 XX Shimkets RA, Leach M;
 XX WPI: 2001-356160/37.
 XX Polymorphic nucleic acid sequences, useful in genetic testing and
 therapy -
 Claim 29; Page 2641; 2653pp; English.
 XX AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
 sequences (I), which contain single nucleotide polymorphisms (SNPs).
 CC AAM53114 to AAM53329 represent peptides related to human polymorphic
 CC polynucleotide sequences. The sequences can be used in gene and protein
 CC therapy, and in vaccine production. (I) and the polypeptides encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate expression of polymorphic polypeptides.
 CC For example, (I) may be used to treat disorders by rectifying mutations
 CC or deletions in a patient's genome that affect the activity of
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patients own production of polypeptide. Additionally, (I) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids
 CC in samples, and therefore which patients may be in need of restorative
 CC therapy. The polypeptides encoded by (I) may be used as antigens in the
 CC production of antibodies specific for polymorphic polypeptides. The
 CC antibodies may also be used to down regulate expression and activity.
 CC The antibodies may also be used as diagnostic agents for detecting the
 CC presence of polymorphic polypeptides in samples.
 XX Sequence 1 AA;
 SQ Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 L 1
 RESULT 14
 AAM53291
 ID AAM53291 standard; Peptide; 1 AA.
 XX AAM53291;
 AC 09-NOV-2001 (first entry)
 XX Human nonconservative amino acid changing SNP related peptide SEQ:6986.
 DE Human; single nucleotide polymorphism; SNP; genome; gene therapy;
 KW protein therapy; vaccine; probe; diagnostic assay; detection;
 KW quantitation; restorative therapy; polymorphic.
 XX Homo sapiens.
 OS WO200140521-A2.
 PN 07-JUN-2001.
 XX 30-NOV-2000; 2000WO-US32758.
 PF 30-NOV-1999; 99US-0168138.
 PR 29-NOV-2000; 2000US-0726173.
 XX (CURA-) CURAGEN CORP.
 XX Shimkets RA, Leach M;
 XX WPI: 2001-356160/37.

XX Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
XX
PS Claim 29; Page 2641; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. NO. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 L 1

• RESULT 15
AAM53328
ID AAM53328 standard; Peptide; 1 AA.
XX
AC AAM53328;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human nonconservative amino acid changing SNP related peptide SEQ:7023.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic.
•
XX Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI; 2001-356160/37.
XX
XX Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
XX
PS Claim 29; Page 2652; 2653pp; English.
XX

CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. NO. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 V 1

Search completed: February 12, 2003, 11:44:37
Job time : 33 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 11:39:59 ; Search time 7.46667 Seconds
(without alignments)
44.439 Million cell updates/sec

Title: US-09-660-302C-1
Perfect score: 8
Sequence: 1 XXXXXXXX 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|------------|--------------------|
| 1 | 0 | 0.0 | 3 | 1 | GRWM_HUMAN | P01157 homo sapien |
| 2 | 0 | 0.0 | 3 | 1 | LUXE_VIBFI | P24272 vibrio fisc |
| 3 | 0 | 0.0 | 3 | 1 | THYL_PIG | P01151 sus scrofa |
| 4 | 0 | 0.0 | 4 | 1 | ACHI_ACHFU | P35904 achatina fu |
| 5 | 0 | 0.0 | 4 | 1 | DCML_PSECH | P19916 pseudomonas |
| 6 | 0 | 0.0 | 4 | 1 | DCMS_PSECH | P19918 pseudomonas |
| 7 | 0 | 0.0 | 4 | 1 | EOSI_HUMAN | P02731 homo sapien |
| 8 | 0 | 0.0 | 4 | 1 | FAR3_HIRME | P42562 hirudo medi |
| 9 | 0 | 0.0 | 4 | 1 | FAR4_HIRME | P42563 hirudo medi |
| 10 | 0 | 0.0 | 4 | 1 | FFKA_ANTEL | P58705 anthopleura |
| 11 | 0 | 0.0 | 4 | 1 | FLRF_HIRME | P42561 hirudo medi |
| 12 | 0 | 0.0 | 4 | 1 | FLRN_ANTEL | P58707 anthopleura |
| 13 | 0 | 0.0 | 4 | 1 | FMRF_MAGNI | P01162 macrocallis |
| 14 | 0 | 0.0 | 4 | 1 | FYRI_ANTEL | P58706 anthopleura |
| 15 | 0 | 0.0 | 4 | 1 | OCPI_OCTMI | P58648 octopus min |
| 16 | 0 | 0.0 | 4 | 1 | OCPI_OCTMI | P58649 octopus min |
| 17 | 0 | 0.0 | 4 | 1 | RM01_YEAST | P36515 saccharomyc |
| 18 | 0 | 0.0 | 4 | 1 | TUFT_HUMAN | P01858 homo sapien |
| 19 | 0 | 0.0 | 5 | 1 | ALL4_CARMA | P81817 carcinus ma |
| 20 | 0 | 0.0 | 5 | 1 | BIOA_CITFR | P13071 citrobacter |
| 21 | 0 | 0.0 | 5 | 1 | BIOB_CITFR | P12997 citrobacter |
| 22 | 0 | 0.0 | 5 | 1 | BPP7_BOTIN | P30425 bothrops in |
| 23 | 0 | 0.0 | 5 | 1 | EI03_LITRU | P82099 litoria rub |
| 24 | 0 | 0.0 | 5 | 1 | EI04_LITRU | P82100 litoria rub |
| 25 | 0 | 0.0 | 5 | 1 | FARP_ARTTR | P41853 artiposthi |
| 26 | 0 | 0.0 | 5 | 1 | PAP2_PARMA | P81864 pardachirus |
| 27 | 0 | 0.0 | 5 | 1 | PRC7_PERAM | P01373 periplaneta |
| 28 | 0 | 0.0 | 5 | 1 | PSK_DAUCA | P58261 daucus caro |
| 29 | 0 | 0.0 | 5 | 1 | RE11_LITRU | P82070 litoria rub |
| 30 | 0 | 0.0 | 5 | 1 | RE11_LITRU | P82071 litoria rub |
| 31 | 0 | 0.0 | 5 | 1 | RE31_LITRU | P82072 litoria rub |
| 32 | 0 | 0.0 | 5 | 1 | RE32_LITRU | P82073 litoria rub |
| 33 | 0 | 0.0 | 5 | 1 | SUGA_ACHDO | P19991 acheta dome |

34 0 0.0 5 1 TPIS_CANFA P54714 canis famli
35 0 0.0 5 1 TRM3_ECOLI P13973 escherichia
36 0 0.0 5 1 UC22_MAIZE P80628 zea mays (m
37 0 0.0 5 1 UF01_MOUSE P38639 mus musculu
38 0 0.0 5 1 UXAA_CHLTR P38005 chlamydia t
39 0 0.0 6 1 ACPH_RABIT P25134 oryctolagus
40 0 0.0 6 1 ASP2_LACSN P82655 lactobacill
41 0 0.0 6 1 CIP1_MYTED P13736 mytilus edu
42 0 0.0 6 1 CIP2_MYTED P13737 mytilus edu
43 0 0.0 6 1 EI01_LITRU P82096 litoria rub
44 0 0.0 6 1 FARP_MONEK P41966 moniezia ex
45 0 0.0 6 1 LOKI_LOCMI P41491 locusta mig

ALIGNMENTS

RESULT 1

GRWM_HUMAN GRWM_HUMAN STANDARD; PRT; 3 AA.
ID P01157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Growth-modulating peptide.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77162369; PubMed=858356;
RA Schlesinger D.H., Pickart L., Thaler M.M.;
RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";
RL Experientia 33:324-325(1977).
CC -!- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE
CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.
DR PIR; A01421; GKHU.
SQ SEQUENCE 3 AA; 340 MW; 6331EB1000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 3;
Best local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 G 1

RESULT 2

LUXE_VIBFI LUXE_VIBFI STANDARD; PRT; 3 AA.
ID P24272;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl-
DE protein synthetase) (fragment).
GN LUXE.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OC NCBI_TaxID=668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91072226; PubMed=2254256;
RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination
RT site for the lux operon.";
RL J. Bacteriol. 172:6797-6802(1990).
CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.

CC -I- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
 CC an acyl-protein thiolester.
 CC -I- PATHWAY: Bioluminescent fatty acid reduction system; second step.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M62812; -; NOT_ANNOTATED_CDS.
 CC DR Bioluminescence; Ligase.
 CC KW NON_TER 1 1
 CC FT SEQUENCE 3 AA; 374 MW; 6AA3303000000000000 CRC64;
 CC SQ
 CC
 CC Query Match 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity 0.0%; Pred. No. 0;
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 1 X 1
 CC 1 I 1
 CC
 CC RESULT 3
 CC THYL_PIG STANDARD; PRT; 3 AA.
 CC ID THYL_PIG
 CC AC P01151;
 CC DT 21-JUL-1986 (Rel. 01, Created)
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
 CC OS Sus scrofa (Pig),
 CC OS Ovis aries (Sheep),
 CC OS Bombina orientalis (Oriental fire-bellied toad), and
 CC OS Bombina orientalis viridescens (Eastern newt) (Triturus viridescens).
 CC OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC OX NCBI_TaxID=9823, 9940, 8346, 8316;
 CC [1]
 CC RN SEQUENCE.
 CC RC SPECIES=Pig; TISSUE=Hypothalamus;
 CC RX MEDLINE=70136150; PubMed=4984938;
 CC RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
 CC RT "Structure of porcine thyrotropin releasing hormone."
 CC RL Biochemistry 9:1103-1106(1970).
 CC RN [2]
 CC RN SYNTHESIS.
 CC RC SPECIES=Pig;
 CC RX MEDLINE=70039904; PubMed=4982117;
 CC RA Boier J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
 CC RT "The identity of chemical and hormonal properties of the thyrotropin
 CC releasing hormone and pyroglutamyl-histidyl-proline amide."
 CC RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
 CC RN [3]
 CC RN SEQUENCE.
 CC RC SPECIES=Sheep; TISSUE=Hypothalamus;
 CC RA Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,
 CC RA Ward D.N.;
 CC RT "The elucidation of the primary structure of the hypothalamic thyroid
 CC stimulating hormone releasing factor of ovine origin by means of mass
 CC spectrometry."
 CC RL Org. Mass Spectrom. 5:221-228(1971).
 CC RN [4]
 CC RN SYNTHESIS.
 CC RC SPECIES=Sheep;
 CC RX MEDLINE=70163386; PubMed=4985794;
 CC RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
 CC RA Guillemin R.;
 CC RT "Characterization of ovine hypothalamic hypophysiotropic
 CC TSH-releasing factor."

RL Nature 226:321-325(1970).
 RN [5]
 RN SEQUENCE.
 RC SPECIES=B.orientalis; TISSUE=Skin;
 RX MEDLINE=76138399; PubMed=815011;
 RA Yasuhara T., Nakajima T.;
 RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin."
 RL Chem. Pharm. Bull. 23:3301-3303(1975).
 RN [6]
 RN SEQUENCE.
 RC SPECIES=N.viridescens;
 RX MEDLINE=75035605; PubMed=4214528;
 RA Grimm-Joergensen Y., McKelvy J.F.;
 RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
 viridescens) brain in vitro. Isolation and characterization of
 thyrotropin releasing factor."
 RL J. Neurochem. 23:471-478(1974).
 CC -I- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
 CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
 CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
 CC
 CC PIR; A01415; RHPGT.
 CC PIR; A93750; RSHFT.
 CC PIR; A90919; RHTDFO.
 CC PIR; A92971; A92971.
 CC KW Amidation.
 CC FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 3 3 AMIDATION.
 CC SQ SEQUENCE 3 AA; 380 MW; 7761F6B0000000000000 CRC64;
 CC
 CC Query Match 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity 0.0%; Pred. No. 0;
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 1 X 1
 CC 1 Q 1
 CC
 CC RESULT 4
 CC ACHL_ACHFU STANDARD; PRT; 4 AA.
 CC ID ACHL_ACHFU
 CC AC P35904;
 CC DT 01-JUN-1994 (Rel. 29, Created)
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE Achatin-I.
 CC OS Achatina fulica (Giant African snail).
 CC OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 CC OC Achatinacea; Achatinidae; Achatina.
 CC OX NCBI_TaxID=6530;
 CC [1]
 CC RN SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 CC RC STRAIN=Ferussac; TISSUE=Ganglion;
 CC RX MEDLINE=89273551; PubMed=2597281;
 CC RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
 CC RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
 CC RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
 CC RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
 CC fulica ferussac containing a D-amino acid residue."
 CC RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 CC RN [2]
 CC RN CHARACTERIZATION.
 CC RC STRAIN=Ferussac; TISSUE=Heart atrium;
 CC RX MEDLINE=91264856; PubMed=1675568;
 CC RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
 CC RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
 CC RT "Purification of achatin-I from the atria of the African giant snail,
 CC Achatina fulica, and its possible function."
 CC RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 CC RN [3]
 CC RN X-RAY CRYSTALLOGRAPHY.
 CC RX MEDLINE=93014529; PubMed=1399265;
 CC RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

```

RA Iwashita T., Nomoto K.;
RT "Crystal structure and molecular conformation of achatin-I
RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
RT D-amino acid residue.";
RL Int. J. Pept. Protein Res. 39:258-264(1992).
CC -|- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.
DR PIR; A32480; A32480.
KW Hormone; D-amino acid.
FT MOD_RES 2 2
SQ SEQUENCE 4 AA; 408 MW; 6AADB9C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 3 A 3

RESULT 5
DCML_PSECH STANDARD; PRT; 4 AA.
AC P19916;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit L) (CO-DH L) (Fragment).
GN CUTL.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group.
OX NCBI_TaxID=290;
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydrotrophic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -|- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -|- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -|- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
CC -|- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P10146; P10146.
KW Oxidoreductase; Iron-sulfur.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 A 2

RESULT 7
BOSI_HUMAN STANDARD; PRT; 4 AA.
ID E051_HUMAN
AC P02731;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilolactac peptides.
DE Eosinophilolactac peptides.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76078412; PubMed=1060093;
RA Goetzl E.J., Austen K.F.;
RT "Purification and synthesis of eosinophilolactac tetrapeptides of
RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -|- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
DR PIR; A03190; ETHUL.
FT VARIANT 1 1 V -> A (IN OTHER PEPTIDE).
FT /FTid=VAR.005201.
SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 3 S 3

RESULT 6
DCMS_PSECH STANDARD; PRT; 4 AA.
ID DCMS_PSECH
AC P19918;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO

```

```

DE dehydrogenase subunit S) (CO-DH S) (Fragment).
GN CUTS.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group.
OX NCBI_TaxID=290;
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydrotrophic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -|- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -|- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -|- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
CC -|- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P10146; P10146.
KW Oxidoreductase; Iron-sulfur.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 A 2

RESULT 7
BOSI_HUMAN STANDARD; PRT; 4 AA.
ID E051_HUMAN
AC P02731;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilolactac peptides.
DE Eosinophilolactac peptides.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76078412; PubMed=1060093;
RA Goetzl E.J., Austen K.F.;
RT "Purification and synthesis of eosinophilolactac tetrapeptides of
RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -|- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
DR PIR; A03190; ETHUL.
FT VARIANT 1 1 V -> A (IN OTHER PEPTIDE).
FT /FTid=VAR.005201.
SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 3 S 3

```

```

RESULT 8
FAR3_HIRME          STANDARD;          PRT;          4 AA.
ID FAR3_HIRME
AC P42562;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide YLRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 598 MW; 69D4073B300000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 Y 1

RESULT 9
FAR4_HIRME          STANDARD;          PRT;          4 AA.
ID FAR4_HIRME
AC P42563;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide YMRP-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 616 MW; 69D4068B300000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 Y 1

RESULT 10
FFKA_ATEL          STANDARD;          PRT;          4 AA.
ID FFKA_ATEL
AC P58705;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Antho-KAamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthaeae; Actiniidae; Anthopleura.
NCBI_TaxID=6110;
RN [1]
RP SEQUENCE.
RX MEDLINE=92028852; PubMed=1681803;
RA Nothacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenyllactyl-L-Phe-Lys-Ala-NH2 (Antho-KAamide), a
novel neuropeptide from sea anemones.";
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN [2]
RP FUNCTION.
RX MEDLINE=93391436; PubMed=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
inhibitory neuropeptides, Antho-KAamide and Antho-Riamide.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle
groups. May be involved in the expansion phase of feeding
behaviour in sea anemones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron-specific.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 L-3-PHENYLLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 512 MW; 6DD339CA000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 4 A 4

RESULT 11
FLRF_HIRME          STANDARD;          PRT;          4 AA.
ID FLRF_HIRME
AC P42561;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FLRFamide.
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
NCBI_TaxID=6421, 27815;
RN [1]
RP SEQUENCE.
RC SPECIES=H.medicalinalis;
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
RN [2]
RP SEQUENCE.
RC SPECIES=H.trivolvis; TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
trivolvis.";
RL Peptides 15:31-36(1994).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 582 MW; 69D40729A00000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;

```

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 F 1

RESULT 12

ID FLRN_ANTEL STANDARD; PRT; 4 AA.
AC P58707;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antho-RNamide.
NS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynantheae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RP SEQUENCE, AND MASS-SPECTROMETRY.
RX MEDLINE=90319122; PubMed=1973541;
RA Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
Reimscheid R.K., Nothacker H.-P., Staley A.L.;
RT "Isolation of L-3-phenyllactyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
anemone neuropeptide containing an unusual amino-terminal blocking
group.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron-specific.
CC -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 L-3-PHENYLLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 549 MW; 64540729A000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 F 1

RESULT 13

ID FMRF_MACNI STANDARD; PRT; 4 AA.
AC P01162;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE FMRFamide (Peak C) (Cardioexcitatory neuropeptide).
OS Macrocallista nimbosa (Sun-ray clam).
OS Nereis virens (Sandworm).
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Veneroidea; Veneridae; Macrocallista.
OX NCBI_TaxID=6594, 6353, 6421, 27815;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC SPECIES=M.nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;
RX MEDLINE=77215956; PubMed=877582;
RA Price D.A., Greenberg M.J.;
RT "Structure of a molluscan cardioexcitatory neuropeptide.";
RL Science 197:670-671(1977).
RN [2]
RP SEQUENCE, AND CHARACTERIZATION.
RC SPECIES=M.nimbosa; TISSUE=Ganglion;
RX MEDLINE=78012038; PubMed=909875;
RA Price D.A., Greenberg M.J.;

RT "Purification and characterization of a cardioexcitatory neuropeptide
from the central ganglia of a bivalve mollusc.";
RL Prep. Biochem. 7:261-281(1977).
RN [3]
RP SEQUENCE.
RC SPECIES=N.virens;
RX MEDLINE=90259866; PubMed=2342992;
RA Krajniak K.G., Price D.A.;
RT "Authentic FMRFamide is present in the polychaete Nereis virens.";
RL Peptides 11:75-77(1990).
RN [4]
RP SEQUENCE.
RC SPECIES=H.medicinalis;
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
RN [5]
RP SEQUENCE.
RC SPECIES=H.trivolvis; TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
trivolvis.";
RL Peptides 15:31-36(1994).

CC -!- FUNCTION: MYOACTIVE; CARDIOEXCITATORY SUBSTANCE. PHARMACOLOGICAL
ACTIVITIES INCLUDE AUGMENTATION, INDUCTION, AND REGULARIZATION OF
CARDIAC CONTRACTION.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
FAMILY.
CC PIR: A01426; ECNK.
DR PIR: A60418; A60418.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 600 MW; 69D40699A000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 F 1

RESULT 14

ID FYRI_ANTEL STANDARD; PRT; 4 AA.
AC P58706;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antho-Riamide I (Contains: Antho-Riamide II).
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynantheae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RP SEQUENCE.
RX MEDLINE=92270459; PubMed=1821096;
RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
RA Grimmelikhuijzen C.J.P.;
RT "Isolation of two novel neuropeptides from sea anemones: the unusual,
biologically active L-3-phenyllactyl-Tyr-Arg-Ile-NH2 and its
des-phenyllactyl fragment Tyr-Arg-Ile-NH2.";
RL Peptides 12:1165-1173(1991).
RN [2]
RP FUNCTION.
RX MEDLINE=933991436; PubMed=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
inhibitory neuropeptides, Antho-KAamide and Antho-Riamide.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).

CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron-specific.
DR InterPro: IPR001023; Hsp70.
KW Neuropeptide; Amidation.
FT CHAIN 1 4 ANTHO-RIAMIDE I.
FT CHAIN 2 4 ANTHO-RIAMIDE II.
FT MOD_RES 1 1 L-3-PHENYLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 F 1

RESULT 15
P1_OCTMI STANDARD; PRT; 4 AA.
AC PS8648;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cardioactive peptides Ocp-1/Ocp-2.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OX Incirrata; Octopodidae; Octopus.
RN NCBI_TaxID=89766;
[1]
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RA MEDLINE=20336815; PubMed=10876044;
RA Iwakoshi E., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor.";
RL Peptides 21:623-630(2000).
CC -!- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. Ocp-2 is a 1000 time less
CC active than Ocp-1.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: Ocp-2 has L-Phe instead of D-Phe.
CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD_RES 2 2 D-PHENYLALANINE.
SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 G 1

Search completed: February 12, 2003, 11:44:58
Job time : 8.46667 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 11:42:45 ; Search time 60.2667 Seconds
(without alignments)
27.351 Million cell updates/sec

Title: US-09-660-302C-1
Perfect score: 8
Sequence: 1 XXXXXXXX 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | Match | Length | ID | Description |
|------------|-------|---------|-------|--------|-----------|--------------------|
| 1 | 0 | 0.0 | 0.0 | 4 | 11 Q08433 | Q08433 rattus norv |
| 2 | 0 | 0.0 | 0.0 | 5 | 2 P83073 | P83073 bacillus ce |
| 3 | 0 | 0.0 | 0.0 | 5 | 10 Q99007 | Q99007 hordeum vul |
| 4 | 0 | 0.0 | 0.0 | 5 | 13 P83308 | P83308 gallus gall |
| 5 | 0 | 0.0 | 0.0 | 6 | 10 P82181 | P82181 spinacia ol |
| 6 | 0 | 0.0 | 0.0 | 6 | 10 P82541 | P82541 spinacia ol |
| 7 | 0 | 0.0 | 0.0 | 6 | 10 P82182 | P82182 spinacia ol |
| 8 | 0 | 0.0 | 0.0 | 7 | 2 O07354 | O07354 synechococc |
| 9 | 0 | 0.0 | 0.0 | 7 | 2 Q47029 | Q47029 enterobacte |
| 10 | 0 | 0.0 | 0.0 | 7 | 2 Q50556 | Q50556 actinobacil |
| 11 | 0 | 0.0 | 0.0 | 7 | 2 Q34028 | Q34028 sphingomona |
| 12 | 0 | 0.0 | 0.0 | 7 | 2 Q47477 | Q47477 escherichia |
| 13 | 0 | 0.0 | 0.0 | 7 | 2 Q47505 | Q47505 escherichia |
| 14 | 0 | 0.0 | 0.0 | 7 | 2 P70804 | P70804 azotobacter |
| 15 | 0 | 0.0 | 0.0 | 7 | 2 Q54248 | Q54248 streptomyce |
| 16 | 0 | 0.0 | 0.0 | 7 | 2 P72081 | P72081 nocardia la |

| | | | | | | |
|----|---|-----|---|----|--------|--------------------|
| 17 | 0 | 0.0 | 7 | 4 | Q15897 | Q15897 homo sapien |
| 18 | 0 | 0.0 | 7 | 4 | Q15903 | Q15903 homo sapien |
| 19 | 0 | 0.0 | 7 | 5 | P83274 | P83274 macrobrachi |
| 20 | 0 | 0.0 | 7 | 6 | Q28742 | Q28742 oryctolagus |
| 21 | 0 | 0.0 | 7 | 8 | P92372 | P92372 haynaldia v |
| 22 | 0 | 0.0 | 7 | 8 | P92403 | P92403 lophopyrum |
| 23 | 0 | 0.0 | 7 | 8 | P92425 | P92425 pseudoroegn |
| 24 | 0 | 0.0 | 7 | 8 | P92387 | P92387 henrardia p |
| 25 | 0 | 0.0 | 7 | 8 | P92427 | P92427 peridictyon |
| 26 | 0 | 0.0 | 7 | 8 | P92390 | P92390 heteranthel |
| 27 | 0 | 0.0 | 7 | 8 | P92226 | P92226 crithopsis |
| 28 | 0 | 0.0 | 7 | 8 | P92214 | P92214 amblyopyrum |
| 29 | 0 | 0.0 | 7 | 8 | P92430 | P92430 aegilops ta |
| 30 | 0 | 0.0 | 7 | 8 | P92221 | P92221 bromus iner |
| 31 | 0 | 0.0 | 7 | 8 | P92442 | P92442 taeniathehu |
| 32 | 0 | 0.0 | 7 | 8 | P92381 | P92381 hordeum bra |
| 33 | 0 | 0.0 | 7 | 8 | P92393 | P92393 hordeum vul |
| 34 | 0 | 0.0 | 7 | 8 | P92218 | P92218 australopyr |
| 35 | 0 | 0.0 | 7 | 8 | P92440 | P92440 thinopyrum |
| 36 | 0 | 0.0 | 7 | 8 | P92210 | P92210 agropyron c |
| 37 | 0 | 0.0 | 7 | 8 | O99182 | O99182 gnatholebia |
| 38 | 0 | 0.0 | 7 | 8 | O95945 | O95945 saccharomyc |
| 39 | 0 | 0.0 | 7 | 8 | O98866 | O98866 spinacia ol |
| 40 | 0 | 0.0 | 7 | 8 | P92421 | P92421 psathyrosta |
| 41 | 0 | 0.0 | 7 | 8 | P92385 | P92385 hordeum mar |
| 42 | 0 | 0.0 | 7 | 10 | O49223 | O49223 glycine max |
| 43 | 0 | 0.0 | 7 | 10 | O9C5B3 | O9C5B3 arabidopsis |
| 44 | 0 | 0.0 | 7 | 10 | P93233 | P93233 lycopersico |
| 45 | 0 | 0.0 | 7 | 10 | P82445 | P82445 nicotiana t |

ALIGNMENTS

RESULT 1

Q08433 PRELIMINARY: PRT; 4 AA.
ID Q08433
AC Q08433;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-JAN-1999 (TReMBLrel. 09, Last annotation update)
DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGT)
(Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUNN;
RX MEDLINE=91282758; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koizumi O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RL hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
CC -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
ENDOGENOUS COMPOUNDS.
CC -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
BETA-D-GLUCURONOSIDE.
CC -!- SUBCELLULAR LOCATION: MICROsome.
DR EMBL; S38636; AAB19259.1; -;
KW Transferase; Glycosyltransferase; Microsome; Multigene family.
FT NON_TER 1 1
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 473 MW; 633732c420000000 CRC64;

Query Match 0.0%; Score 0; DB 11; Length 4;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 x 1

Db 1 N 1

RESULT 2

ID P83073 PRELIMINARY; PRT; 5 AA.
 AC P83073;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE 88 kDa protein (Fragment).
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=NCIMB 11796;
 RA Browne N., Dowds B.C.A.;
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.
 FT NON_TER 5 5
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F000000 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 4 T 4

RESULT 3

ID Q99007 PRELIMINARY; PRT; 5 AA.
 AC Q99007;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Alpha-amylase (EC 3.2.1.1) (Fragment).
 GN Aml1.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HIMALAYA; TISSUE=ALEURONE LAYER;
 RX MEDLINE=91329704; PubMed=1831055;
 Y Jacobsen J.V., Close T.J.;
 "Control of transient expression of chimaeric genes by gibberellic acid and abscisic acid in protoplasts prepared from mature barley aleurone layers.";
 RL Plant Mol. Biol. 16:713-721(1991).
 CC -1- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA-AMYLASE IN BARLEY.
 DR EMBL; X54643; CAA38455.1;
 DR Hydrolyase; Glycosidase; Carbohydrate metabolism; Seed; Germination;
 KW Calcium; Multigene family.
 FT NON_TER 5 5
 SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F000000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 2 A 2

RESULT 4

ID P83308 PRELIMINARY; PRT; 5 AA.
 AC P83308;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE FMRFamide-like neuropeptide (LPLRF-amide).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=BRAIN;
 RX PubMed=6137771;
 RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
 RT "A novel active pentapeptide from chicken brain identified by antibodies to FMRFamide.";
 RL Nature 305:328-330(1983).
 CC -1- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.
 CC Neuropeptide.
 KW SEQUENCE 5 AA; 645 MW; 69D40737674000000 CRC64;

Query Match 0.0%; Score 0; DB 13; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 L 1

RESULT 5

ID P82181 PRELIMINARY; PRT; 6 AA.
 AC P82181;
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. ALVARO; TISSUE=LEAF;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamaguchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast).";
 RL J. Biol. Chem. 275:28466-28482(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro; IPR001790; Ribosomal_L10.
 DR ProFam; PF00466; Ribosomal_L10; PARTIAL.
 DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON_TER 6 6
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB0000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;


```

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 A 1

RESULT 6
P82541 ID P82541 PRELIMINARY; PRT; 6 AA.
AC P82541;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC STRAIN=CV. ALVARO; TISSUE=LEAF;
RX MEDLINE=20435797; PubMed=10874039;
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RL J. Biol. Chem. 37:28455-28465(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
CC FORM IS THE MINOR BASIC FORM.
CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
CC InterPro: IPR002222; Ribosomal_S19.
DR Pfam: PF00203; Ribosomal_S19; PARTIAL.
DR PRINTS: PR00975; RIBOSOMALS19; PARTIAL.
DR PROSITE: PS00323; RIBOSOMAL_S19; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 T 1

RESULT 7
P82182 ID P82182 PRELIMINARY; PRT; 6 AA.
AC P82182;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALVARO; TISSUE=LEAF;
RX MEDLINE=20435798; PubMed=10874046;

```

```

Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RL J. Biol. Chem. 275:28466-28482(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00466; Ribosomal_L10; PARTIAL.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 A 1

RESULT 8
O07354 ID O07354 PRELIMINARY; PRT; 7 AA.
AC O07354;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NifK (Fragment).
GN NIFK.
OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
OX NCBI_TaxID=41431;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RF-1;
RX MEDLINE=99231861; PubMed=10217509;
RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT "Organization and expression of nitrogen-fixation genes in the aerobic
RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT RF-1."
RL Microbiology 145:743-753(1999).
DR EMBL: AF003700; AAC35193.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 S 2

RESULT 9
Q47029 ID Q47029 PRELIMINARY; PRT; 7 AA.
AC Q47029;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Aad A1 protein (Fragment).
GN Aad A1.
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=550;

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=94079349; PubMed=8257126;
RA Ratner P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT "Analysis of the aac(3)-Via gene encoding a novel 3-N-
RT acetyltransferase."
RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR EMBL; M88012; AAA16193.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 2 T 2

RESULT 10
O50556 PRELIMINARY; PRT; 7 AA.
O50556;
O1-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE GlyA (Fragment).
GN GlyA.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33384;
RX MEDLINE=96355846; PubMed=8751884;
RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,
RA Kraig E.;
RT "cis Elements and trans factors are both important in strain-specific
RT regulation of the leukotoxin gene in Actinobacillus
RT actinomycetemcomitans."
RL Infect. Immun. 64:3451-3460(1996).
DR EMBL; U51862; AAB88721.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 X 1
7 A 7

Db

RESULT 11
O34028 PRELIMINARY; PRT; 7 AA.
AC O34028;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE Catechol-2,3-dioxygenase (Fragment).
GN PHNE.
OS Sphingomonas chungbukensis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC Sphingomonas.
OX NCBI_TaxID=56193;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DJ77;

```

```

RA Kim Y.-C.;
RP Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88298; AAB66311.1; -.
KW Dioxygenase. 1 1
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 868 MW; 71A452D1A699D460 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 4 T 4

RESULT 12
Q47477 PRELIMINARY; PRT; 7 AA.
AC Q47477;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Tpi protein (Fragment).
GN TPI.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE OF 7-7 FROM N.A.
RX MEDLINE=85203917; PubMed=3158524;
RA Hellinda H.W., Evans P.R.;
RT "Nucleotide sequence and high-level expression of the major
RT Escherichia coli phosphofructokinase."
RL Eur. J. Biochem. 149:363-373(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Evans P.;
RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
DR EMBL; X02519; CAA26359.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 773 MW; 7416D33DDDB1DB0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 A 1

RESULT 13
Q47505 PRELIMINARY; PRT; 7 AA.
AC Q47505;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MccA protein.
GN MCCA.
OS Escherichia coli.
OG Plasmid pMccC7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96099297; PubMed=8522520;
RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;
RT "Structure and organization of plasmid genes required to produce the
RT translation inhibitor microcin C7."

```

RL J. Bacteriol. 177:7131-7140(1995).
DR EMBL; X57583; CAA40808.1; -.
KW Plasmid.
SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 3 T 3

RESULT 14

P70804 PRELIMINARY; PRT; 7 AA.
ID P70804;
IC P70804;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Algt protein (Fragment).
GN ALGT.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-E;
RX MEDLINE=96427318; PubMed=8830682;
RA Rehm B.H.A., Ertesvag H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87973; CAA61230.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 2 T 2

RESULT 15

Q54248 PRELIMINARY; PRT; 7 AA.
ID Q54248;
AC Q54248;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RplO protein (Fragment).
GN RplO.
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-N2-3-11;
RX MEDLINE=20011291; PubMed=10542330;
RA Poehling S., Piepersberg W., Wehmeier U.F.;
RT "Analysis and regulation of the sec Y gene from Streptomyces griseus
RT N2-3-11 and interaction of the SecY protein with the SecA protein.";
RL Biochim. Biophys. Acta 1447:298-302(1999).
DR EMBL; X95915; CAA65160.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 2 T 2

Search completed: February 12, 2003, 11:46:58
Job time : 61.2667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 11:39:30 ; Search time 28 seconds
(without alignments)
33.313 Million cell updates/sec

Title: US-09-660-302C-7
Perfect score: 43
Sequence: 1 CEEDFYR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|-----------------------|
| 1 | 43 | 100.0 | 7 | AA1980 | Proteolytic cleavage |
| 2 | 38 | 88.4 | 33 | AA1981 | Plasmodium falciparum |
| 3 | 38 | 88.4 | 246 | AA1982 | Human growth hormone |
| 4 | 38 | 88.4 | 249 | AA1983 | Hormone binding re |
| 5 | 38 | 88.4 | 269 | AA1984 | Soluble human grow |
| 6 | 38 | 88.4 | 269 | AA1985 | Human somatogenic |
| 7 | 38 | 88.4 | 269 | AA1986 | Human soluble grow |
| 8 | 38 | 88.4 | 269 | AA1987 | Human soluble soma |
| 9 | 38 | 88.4 | 269 | AA1988 | Soluble part of th |
| 10 | 38 | 88.4 | 315 | AA1989 | Human fusion prote |

| | | | | | | |
|----|----|------|------|----|----------|--------------------|
| 11 | 38 | 88.4 | 340 | 23 | AAU75496 | Human fusion prote |
| 12 | 38 | 88.4 | 637 | 10 | AAU75496 | Human growth hormo |
| 13 | 38 | 88.4 | 638 | 9 | AAU75496 | Human growth hormo |
| 14 | 38 | 88.4 | 638 | 19 | AAU75496 | Human growth hormo |
| 15 | 38 | 88.4 | 648 | 22 | AAU75496 | Human growth hormo |
| 16 | 36 | 83.7 | 84 | 20 | AAU75496 | Mature nematode ex |
| 17 | 36 | 83.7 | 84 | 21 | AAU75496 | A. caninum nematod |
| 18 | 36 | 83.7 | 91 | 17 | AAU75496 | AcanaPC2, Ancylos |
| 19 | 36 | 83.7 | 91 | 20 | AAU75496 | Nematode extracted |
| 20 | 36 | 83.7 | 91 | 20 | AAU75496 | Nematode extracted |
| 21 | 36 | 83.7 | 91 | 21 | AAU75496 | A. caninum nematod |
| 22 | 36 | 83.7 | 404 | 19 | AAU75496 | LH-2 protein #1. |
| 23 | 36 | 83.7 | 406 | 22 | AAU75496 | Human protein sequ |
| 24 | 36 | 83.7 | 423 | 19 | AAU75496 | LH-2 protein #2. |
| 25 | 34 | 79.1 | 49 | 22 | AAU75496 | Human polypeptide |
| 26 | 34 | 79.1 | 102 | 23 | AAU75496 | Streptococcus poly |
| 27 | 34 | 79.1 | 615 | 22 | AAU75496 | Novel human diagno |
| 28 | 34 | 79.1 | 1589 | 22 | AAU75496 | Interferon induced |
| 29 | 34 | 79.1 | 1591 | 22 | AAU75496 | Peptide #4764 enco |
| 30 | 34 | 79.1 | 1591 | 22 | AAU75496 | Peptide #4871 enco |
| 31 | 34 | 79.1 | 1591 | 22 | AAU75496 | Protein #4653 enco |
| 32 | 34 | 79.1 | 1591 | 22 | AAU75496 | Human brain expres |
| 33 | 34 | 79.1 | 1591 | 22 | AAU75496 | Human bone marrow |
| 34 | 34 | 79.1 | 1591 | 22 | AAU75496 | Peptide #4765 enco |
| 35 | 34 | 79.1 | 1591 | 22 | AAU75496 | Peptide #4856 enco |
| 36 | 34 | 79.1 | 1591 | 22 | AAU75496 | Peptide #4620 enco |
| 37 | 34 | 79.1 | 1591 | 23 | AAU75496 | Human peptide enco |
| 38 | 33 | 76.7 | 75 | 20 | AAU75496 | Mature nematode ex |
| 39 | 33 | 76.7 | 77 | 20 | AAU75496 | Mature nematode ex |
| 40 | 33 | 76.7 | 78 | 20 | AAU75496 | Mature nematode ex |
| 41 | 33 | 76.7 | 78 | 21 | AAU75496 | A. caninum nematod |
| 42 | 33 | 76.7 | 82 | 20 | AAU75496 | Nematode extracted |
| 43 | 33 | 76.7 | 82 | 20 | AAU75496 | Mature nematode ex |
| 44 | 33 | 76.7 | 82 | 21 | AAU75496 | A. ceylanicum nema |
| 45 | 33 | 76.7 | 82 | 21 | AAU75496 | A. ceylanicum nema |

ALIGNMENTS

RESULT 1
AAU75496
ID AAU75496 standard; peptide; 7 AA.
AC AAU75496;
XX 09-NOV-1999 (first entry)
DT Proteolytic cleavage signal site used in inhibiting receptor proteolysis.
DE Signal transduction; proteolytic cleavage; cleavage signal site;
KW ubiquitin; proteosome binding site; muscle wasting; renal tubular defect;
KW uremia; diabetes; Cushing's disease; eating disorder; AIDS;
KW growth hormone deficiency.
OS Mammalia.
XX EP943624-A1.
PN 22-SEP-1999.
PD 12-MAR-1998; 98EP-0200799.
PF 12-MAR-1998; 98EP-0200799.
PR (UYUT-) RIJKSUNIV UTRECHT.
PA WPI; 1999-510568/43.
DR Controlling the availability and/or signal transduction capability
XX of a cell surface receptor, useful for treating growth hormone
PT deficiencies

PS Claim 15; Page 27; 36pp; English.

XX This sequence is a cleavage signal site. This site is used in a method
CC for controlling the availability and signal transduction capability of a
CC cell surface receptor by administering an inhibitor that is capable of
CC inhibiting proteolytic cleavage of the receptor. Inhibition of this
CC proteolytic cleavage results in the receptors being present on the
CC surface for longer and therefore signalling for longer to the interior of
CC the cell. This increases the sensitivity of cells to any hormones which
CC might be present. The inhibitor is derived from or competes with an
CC amino acid sequence around this proteolytic cleavage signal. The
CC inhibitor may be used to treat muscle wasting, associated with disorders
CC such as renal tubular defects, uraemia, diabetes, Cushing's syndrome,
CC cachexias, eating disorders, AIDS, after stress and during neuromuscular
CC disease.

XX
XX
XX Sequence 7 AA;

Query Match 100.0%; Score 43; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
| | | | |
1 CEEDFYR 7

RESULT 2
AAR22456
ID AAR22456 standard; Protein; 33 AA.
XX
AC AAR22456;
XX
DT 09-SEP-1992 (first entry)
XX
DE Plasmid pBSR8-3 encoded polypeptide.
XX
DE Recombinant; plasmid pACYMhGHR; hGHR binding region;
KW human growth hormone.
KW
XX
OS Synthetic.
OS
XX
XX JP04063594-A.
PN
XX
XX 28-FEB-1992.
PD
XX
XX 03-JUL-1990; 90JP-0176630.
PF
XX
XX 03-JUL-1990; 90JP-0176630.
PR
XX
XX (TANP-) TANPAKU KAGAKU KENK.
PA
XX
XX WPI; 1992-120154/15.
DR N-PSDB; AAQ23311.
XX
XX New recombinant plasmid and Baculovirus having specific DNA
PT fragment - used for prep. of hormone binding region protein of
PT human growth hormone receptor
PT
XX
PS Example; Fig 2; 11pp; Japanese.
PS
XX
XX The sequence is that encoded by a fragment of the plasmid pBSR8-3
CC which is used in the construction of plasmid pACYMhGHR. pACYMhGHR
CC contains a DNA fragment encoding the binding region of human growth
CC hormone receptor (hGHR) under control of a Baculovirus gene expression
CC system. This can be used to transform insect cells to produce a
CC more natural protein than is obtd. with E.coli and in larger amts.
CC See also AAR22457-R22459.

XX
XX
XX Sequence 33 AA;

Query Match 88.4%; Score 38; DB 13; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.9;

QY 1 CEEDFY 6
| | | | |
28 CEEDFY 33

RESULT 3
AAR56389
ID AAR56389 standard; protein; 246 AA.
XX
AC AAR56389;
XX
DT 01-MAR-1995 (first entry)
XX
XX Human growth hormone receptor C-domain.
DE
XX
XX Human growth hormone receptor; hGHR; C-domain; CD; hGHR-CD;
KW extracellular; enhancement; Met130Q-238Q; GS129V-238Q.
XX
XX Homo sapiens.
OS
XX
XX JP06172394-A.
PN
XX
XX 21-JUN-1994.
PD
XX
XX 10-DEC-1992; 92JP-0330361.
PF
XX
XX 10-DEC-1992; 92JP-0330361.
PR
XX
XX (TANP-) TANPAKU KOGAKU KENKYUSHO KK.
PA
XX
XX WPI; 1994-238767/29.
DR
XX
XX Extracellular C-domain protein of growth hormone receptor
PT (hGHR-CD) - is used to enhance growth hormone function
PT
XX
XX Disclosure; Page 9-10; 16pp; Japanese.
PS
XX
XX The extracellular C-domain of hGHR is used to enhance growth hormone
CC function. Two forms, Met130Q-238Q (AAQ66550) and GS129V-238Q (AAQ66553)
CC are given. Primers for the isolation of extracellular hGHR-CD are
CC given in AAQ66551-52 and AAQ66554-55.

XX
XX
XX Sequence 246 AA;

Query Match 88.4%; Score 38; DB 15; Length 246;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6
| | | | |
241 CEEDFY 246

RESULT 4
AAR06867
ID AAR06867 standard; protein; 249 AA.
XX
AC AAR06867;
XX
XX 16-JAN-1991 (first entry)
DT
XX
XX Hormone binding region of human growth hormone receptor.
DE
XX
XX Synthetic.
OS
XX
XX hGHR.
KW
XX
XX JP02200186-A.
PN
XX
XX 08-AUG-1990.
PD
XX

PF 30-JAN-1989; 89JP-0020182.
XX
PR 30-JAN-1989; 89JP-0020182.
XX
XX (TANP-) TANPAKU KOGAKU KENK.
XX
DR WPI; 1990-285858/38.
XX
DR N-PSDB; AAQ05968.
XX

Synthetic gene - used for coding the amino acid sequence in the hormone-combining region of human growth hormone receptor

Disclosure; Fig 2; l1pp; Japanese.

The sequence is the same as that of the natural receptor but is encoded by a synthetic gene that has a slightly altered nucleotide sequence incorporating at least 2 new restriction sites and removing a direct repeat and a palindromic sequence.

XX Sequence 249 AA;

Query Match 88.4%; Score 38; DB 11; Length 249;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6
|||||
Db 241 CEEDFY 246

RESULT 5
AAR05045
ID AAR05045 standard; protein; 269 AA.

XX
XX
AC AAR05045;

XX 03-OCT-1990 (first entry)

Soluble human growth hormone receptor.

XX Plasmid pJ1446; human growth hormone; somatogenic receptor; mutagenesis;
XX substitutions; active domain; hormone variants.

XX Synthetic.

XX WO9004788-A.

XX 03-MAY-1990.

XX 30-OCT-1989; 89WO-US04778.

XX 29-OCT-1988; 88US-0264611; WO-U04778.

XX (GETH) GENENTECH INC.

XX Wells JA, Cunningham BC;

XX WPI; 1990-164120/21.

XX N-PSDB; AAQ04672.

XX Identifying active domains and amino acid(s) in polypeptide(s) -
XX by specific substitutions, then testing modified products for
XX interaction with target, and new polypeptide, esp. hormone etc.

XX Disclosure; : 3pp; English.

XX The soluble human growth hormone receptor shGHR was subcloned into
XX pB0475 to form pJ1446. E.coli W3110, depp was transformed with
XX pJ1446 and grown in low-phosphate media in a fermentor at 30 degrees C.
XX This 246 amino acid hGHR is produced.
XX See also AAQ04671 and AAQ04672.

XX Sequence 269 AA;

Query Match 88.4%; Score 38; DB 11; Length 269;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6
|||||
Db 264 CEEDFY 269

RESULT 6
AAW10426
ID AAW10426 standard; Protein; 269 AA.

XX AC AAW10426;

XX 11-AUG-1997 (first entry)

XX Human somatogenic receptor extracellular domain.

XX Active site; active domain; growth hormone; somatogenic receptor;
XX mutagenesis.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..23
FT /label= Sig_peptide
FT Protein 24..269
FT /label= Mat_protein
FT /note= "shGHR(aal-246)";

XX US5580723-A.

XX 03-DEC-1996.

XX 28-OCT-1988; 88US-0264611.

XX 26-OCT-1989; 89US-0428066.

XX 28-OCT-1988; 88US-0264611.

XX 27-APR-1992; 92US-0875204.

XX 13-OCT-1992; 92US-0960227.

XX 02-FEB-1994; 94US-0190723.

XX (GETH) GENENTECH INC.

XX Cunningham BC, Wells JA;

XX WPI; 1997-033563/03.

XX N-PSDB; AAT10426.

XX Identification of unknown active domains in polypeptide(s) - useful
XX for analysis of structure and function of hormones, etc.

XX Example 3; Fig 12A-C; 86pp; English.

XX A polypeptide (AAW10426) comprises the soluble extracellular domain
XX of human liver growth hormone receptor (somatogenic receptor,
XX shGHR). It was expressed in Escherichia coli transformants using
XX a vector derived from pJ1446 (AAT47449). shGHR can be used in a
XX novel method designed for the identification of polypeptide active
XX domains. For human growth hormone (hGH) (see also AAW10425), this
XX involves substituting selected amino acid segments of hGH with
XX analogous segments from analogue polypeptides (human placenta
XX lactogen, human prolactin and pig growth hormone) and examining the
XX effect of the substn. on binding affinity to shGHR. Once active
XX site regions are detd., active site amino residues (see also
XX AAW10427-62) can be similarly identified.

XX Sequence 269 AA;

Query Match 88.4%; Score 38; DB 18; Length 269;
Best Local Similarity 100.0%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6
|||||

Db 264 CEEDFY 269

RESULT 7
AAV31767
ID AAY31767 standard; Protein; 269 AA.

XX AC AAY31767;

XX DT 06-DEC-1999 (first entry)

XX DE Human soluble growth hormone receptor.

XX KW Growth hormone receptor; somatogenic receptor; human; pJ1446;
variant; protein engineering.

XX OS Homo sapiens.

XX FT Key Location/Qualifiers
Peptide 1..23
/note= "signal peptide"
Protein 24..269
/note= "mature protein"

XX PN US5955346-A.

XX PD 21-SEP-1999.

XX PF 07-JUN-1995; 95US-0476999.

XX PR 02-FEB-1994; 94US-0190723.

XX PR 26-OCT-1989; 89US-0428066.

XX PR 27-APR-1992; 92US-0875204.

XX PR 13-OCT-1992; 92US-0960227.

XX PR 28-OCT-1988; 88US-0264611.

XX PA (GETH) GENENTECH INC.

XX PI Cunningham BC, Wells JA;

XX DR WPI: 1999-560495/47.

XX DR N-PSDB; AAX87977.

XX PT Isolated nucleic acids encoding variants of human prolactin and
placental lactogen useful for identifying active domains within those
proteins -

XX PT Example 2; Fig 12A-C; 86pp; English.

XX CC This sequence represents a human soluble growth hormone receptor
(shGHR) encoded by plasmid pJ1446 (see AAX87977). shGHR was
expressed in E. coli and was used in binding assays of human growth
hormone variants. The invention provides a method for the
systematic analysis of the structure and function of polypeptides
by identifying active domains which influence the activity of the
polypeptide with a target substance, and a method for identifying
the active amino acid residues within the active domain of a
polypeptide. It also provides polypeptide variants comprising
segment-substituted and residue-substituted growth hormones.
prolactins (see AAY31764) and placental lactogens (see AAY31765).
Identifying receptor binding sites in hormones permits the rational
design of receptor specific variants.

XX SQ Sequence 269 AA;

Query Match 88.4%; Score 38; DB 20; Length 269;

Best Local Similarity 100.0%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6
|||||

Db 264 CEEDFY 269

RESULT 8
AAW82802
ID AAW82802 standard; Protein; 269 AA.

XX AC AAW82802;

XX DT 01-MAR-1999 (first entry)

XX DE Human soluble somatogenic receptor.

XX KW Somatogenic receptor; growth hormone; human; plasmid pJ1446;
active domain.

XX OS Homo sapiens.

XX FT Key Location/Qualifiers
Peptide 1..23
/label= sig_peptide
Protein 24..269
/label= Mat_protein

XX PN US5834250-A.

XX PD 10-NOV-1998.

XX PF 30-JUN-1997; 97US-0903398.

XX PR 26-OCT-1989; 89US-0428066.

XX PR 28-OCT-1988; 88US-0264611.

XX PR 27-APR-1992; 92US-0875204.

XX PR 13-OCT-1992; 92US-0960227.

XX PR 02-FEB-1994; 94US-0190723.

XX PR 06-JUN-1995; 95US-0483039.

XX PR 30-JUN-1997; 97US-0903398.

XX PA (GETH) GENENTECH INC.

XX PI Cunningham BC, Wells JA;

XX DR WPI: 1999-008714/01.

XX DR N-PSDB; AAV62766.

XX PT Identifying amino acids in polypeptide(s) that are involved in
interaction with target - by introducing scanning amino acid
substitutions at selected positions and screening for any change in
interaction, particularly to engineer hormones with altered
properties

XX PT Example 3; Fig 12A-J; 84pp; English.

XX CC This is the amino acid sequence of human liver soluble somatogenic
receptor, as encoded by a DNA insert of plasmid pJ1446 (see
AAV62766). The invention provides methods for the systematic
analysis of the structure and function of polypeptides by
identifying active domains which influence the activity of the
polypeptide with a target substance (for human growth hormone (see
AAW82801), the target used was somatogenic receptor). Active domains
are determined by substituting selected amino acid segments of the
polypeptide with an analogous polypeptide segment, and comparing
the activity of the substituted polypeptide with that of the native
polypeptide for the target. The invention also provides methods
for identifying the active amino acids within the active domain.
The method is particularly applied to hormones. Polypeptides can
be produced that have better biological, biochemical and
immunogenic properties than wild-type proteins, e.g. human growth
hormone with increased potency but reduced diabetogenic activity
and human prolactin or placental lactogen may have increased
activity at somatogenic receptors.

XX SQ Sequence 269 AA;

Query Match 88.4%; Score 38; DB 20; Length 269;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CEEDFY 6
|||||

Db 264 CEEDFY 269

RESULT 9
AAU78429
ID AAY78429 standard; Protein; 269 AA.

XX AC AAY78429;

XX DT 09-MAY-2000 (first entry)

XX DE Soluble part of the somatogenic receptor encoded by plasmid pJ1446.

XX KW Human growth hormone; hGH; prolactin; placental lactogen;
modification; mutagenesis.

XX KW osteoporosis; diabetes mellitus; cancer; obesity; insulin resistance;
hyperlipidaemia; hypertension; anaemia; autoimmune disease;
infectious disease; inflammatory disorder; rheumatoid arthritis;
interleukin-6 chimaera; IL-6.

XX OS Homo sapiens.

XX PN US6013478-A.

XX PD 11-JAN-2000.

XX PF 24-JUN-1998; 98US-0104036.

XX PR 26-OCT-1989; 89US-0428066.

XX PR 27-APR-1992; 92US-0875204.

XX PR 13-OCT-1992; 92US-0960227.

XX PR 02-FEB-1994; 94US-0190723.

XX PR 06-JUN-1995; 95US-0483039.

XX PR 30-JUN-1997; 97US-0903398.

XX PR 28-OCT-1988; 88US-0264611.

XX PA (GETH) GENENTECH INC.

XX PI Wells JA, Cunningham BC;

XX DR WPI; 2000-159873/14.

XX DR N-PSDB; AAZ88448.

XX PT Recombinant production of variant polypeptides, e.g. growth hormone
variants with altered receptor specificity, using cells transformed
with DNA selected by scanning mutagenesis in at least one peptide
domain

XX PS Example 3; Fig 12; 83pp; English.

XX CC The present invention describes the production of a polypeptide variant
(1) comprising segment substituted and residue substituted growth
hormone, prolactin or placental lactogens. The method is particularly
used to produce variants of growth hormone (GH), prolactin or placental
lactogen, but may also be applied to receptors, interferons, and
colony-stimulating factors. A particular application is the production
of human GH variants with altered (decreased or increased) binding
interaction with the somatogenic receptor, i.e. compounds useful as
human GH (ant)agonists and which may have higher potency for stimulating
other human GH receptors, and as standards or tracers in immunoassays
for human GH. This method of DNA selection identifies the biologically
active residues in active domains, including those critical for
interaction with different targets. The present sequence represents the
soluble region of the somatogenic receptor from liver encoded by a
plasmid, which is used in the exemplification of the present invention.

XX SQ Sequence 269 AA;

Query Match 88.4%; Score 38; DB 21; Length 269;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CEEDFY 6
|||||

Db 264 CEEDFY 269

RESULT 10
AAU75499
ID AAU75499 standard; Protein; 315 AA.

XX AC AAU75499;

XX DT 08-MAY-2002 (first entry)

XX DE Human fusion protein Chi 1A2.

XX KW Human; GHSL-23; human growth hormone; GH; GHR;
growth hormone; receptor; GHstopGHR SD100; GHlinkGHR; GHlinkGHRflec;
pTrcRNSsacI; TrcHindrev; Chi 1A2 chimaera; acromegaly; gigantism;
growth hormone deficiency; Turner's syndrome; renal failure;
osteoporosis; diabetes mellitus; cancer; obesity; insulin resistance;
hyperlipidaemia; hypertension; anaemia; autoimmune disease;
infectious disease; inflammatory disorder; rheumatoid arthritis;
interleukin-6 chimaera; IL-6.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 14 /note= "Encoded by ATG"

XX FT Misc-difference 312 /label= "Unknown"

XX FT Misc-difference 313 /note= "Encoded by TGA, in-frame stop codon"

XX FT Misc-difference 313 /label= "Unknown"

XX FT Misc-difference 313 /note= "Encoded by TAA, in-frame stop codon"

XX PN WO200196565-A2.

XX PD 20-DEC-2001.

XX PF 18-JUN-2001; 2001WO-GB02645.

XX PR 16-JUN-2000; 2000GB-0014765.

XX PR 10-MAR-2001; 2001GB-0005969.

XX PR 16-MAR-2001; 2001GB-0006487.

XX PA (ASTE-) ASTERION LTD.

XX PI Ross R, Artymiuik P, Sayers J;

XX DR WPI; 2002-130734/17.

XX DR N-PSDB; ABK14547.

XX PT New binding agent useful in producing a medicament for treating e.g.
cancer, obesity, acromegaly or gigantism, comprises a first part that
binds to a ligand binding domain of a receptor and a second part having
a receptor binding domain

XX PS Claim 49; Fig 22; 79pp; English.

XX CC The invention relates to a binding agent comprising a first part capable
of binding a ligand binding domain of a receptor linked to a second part
comprising a receptor binding domain, where the binding agent modulates
the activity of the receptor. Also included are a nucleic acid molecule
having a sequence, which encodes a binding agent comprising sequences
given in the specification comprising the sequences of the full length
GHstopGHR SD100 construct, GHlinkGHR construct(GH, growth hormone,
GHR, growth hormone receptor), GHlinkGHRflec construct, 1157 base pair

CC PCR fragment GHlinkGHR generated by nucleotides pTcrNSSsacI and
 CC TrcHndrev, or the nucleotide sequence of the Chi 1A2 chimera,
 CC sequences binding to the nucleic acids or degenerate sequences
 CC representing them (which have receptor antagonising activity),
 CC their encoded polypeptides, a vector comprising the nucleic acids and a
 CC cell transformed/transfected with the nucleic acid or vector.
 CC The binding agent is useful for manufacturing a medicament for
 CC the treatment of acromegaly, gigantism, growth hormone (GH) deficiency,
 CC Turner's syndrome, renal failure, osteoporosis, diabetes mellitus,
 CC cancer, obesity, insulin resistance, hyperlipidaemia, hypertension,
 CC anaemia, autoimmune and infectious diseases, and inflammatory disorders
 CC including rheumatoid arthritis (interleukin (IL)-6 chimera).
 CC The present sequence represents the fusion protein agent of the invention
 CC being the Chi 1A2 construct (growth hormone/growth hormone
 CC receptor not linked by a synthetic peptide linker).
 XX
 XX

SQ Sequence 315 AA;

Query Match 88.4%; Score 38; DB 23; Length 315;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CEEDFY 6
 |||||
 306 CEEDFY 311

RESULT 11

AAU75496
 ID AAU75496 standard; Protein: 340 AA.

AC AAU75496;

DT 08-MAY-2002 (first entry)

DE Human fusion protein GHlinkGHR.

XX Human; GHs1-23; human growth hormone; GH; GHR;
 KW growth hormone; receptor; GHstopGHR SD100; GHlinkGHR; GHlinkGHRflec;
 KW pTcrNSSsacI; TrcHndrev; Chi 1A2 chimera; acromegaly; gigantism;
 KW growth hormone deficiency; Turner's syndrome; renal failure;
 KW osteoporosis; diabetes mellitus; cancer; obesity; insulin resistance;
 KW hyperlipidaemia; hypertension; anaemia; autoimmune disease;
 KW infectious disease; inflammatory disorder; rheumatoid arthritis;
 KW interleukin-6 chimera; IL-6.

XX Homo sapiens.
 OS Synthetic.

Key Location/Qualifiers

Misc-difference 14 /note= "Encoded by ATG"

Misc-difference 337 /label= Unknown

FT /note= "Encoded by TGA, in-frame stop codon"

FT Misc-difference 338 /label= Unknown

FT /note= "Encoded by TAA, in-frame stop codon"

XX WO200196565-A2.

PN 20-DEC-2001.

XX 18-JUN-2001; 2001WO-GB02645.

XX 16-JUN-2000; 2000GB-0014765.

PR 10-MAR-2001; 2001GB-0005969.

PR 16-MAR-2001; 2001GB-0006487.

XX (ASTE-) ASTERION LTD.

PA Ross R, Artymuk P, Sayers J;

XX

XX

XX

XX

XX

XX

DR WPI; 2002-130734/17.
 DR N-PSDB; ABK14531.

XX New binding agent useful in producing a medicament for treating e.g.
 PT cancer, obesity, acromegaly or gigantism, comprises a first part that
 PT binds to a ligand binding domain of a receptor and a second part having
 PT a receptor binding domain -

XX Disclosure; Fig 6; 79pp; English.

XX The invention relates to a binding agent comprising a first part capable
 CC of binding a ligand binding domain of a receptor linked to a second part
 CC comprising a receptor binding domain, where the binding agent modulates
 CC the activity of the receptor. Also included are a nucleic acid molecule
 CC having a sequence, which encodes a binding agent comprising sequences
 CC given in the specification comprising the sequences of the full length
 CC GHstopGHR SD100 construct, GHlinkGHR construct(GH, growth hormone,
 CC GHR, growth hormone receptor), GHlinkGHRflec construct, 1157 base pair
 CC PCR fragment GHlinkGHR generated by nucleotides pTcrNSSsacI and
 CC TrcHndrev, or the nucleotide sequence of the Chi 1A2 chimera,
 CC sequences binding to the nucleic acids or degenerate sequences
 CC representing them (which have receptor antagonising activity),
 CC their encoded polypeptides, a vector comprising the nucleic acids and a
 CC cell transformed/transfected with the nucleic acid or vector.
 CC The binding agent is useful for manufacturing a medicament for
 CC the treatment of acromegaly, gigantism, growth hormone (GH) deficiency,
 CC cancer, obesity, insulin resistance, hyperlipidaemia, hypertension,
 CC anaemia, autoimmune and infectious diseases, and inflammatory disorders
 CC including rheumatoid arthritis (interleukin (IL)-6 chimera).
 CC The present sequence represents the fusion protein agent of the invention
 CC being the GHlinkGHR construct (growth hormone/growth hormone
 CC receptor linked by a synthetic peptide linker).
 XX

SQ Sequence 340 AA;

Query Match 88.4%; Score 38; DB 23; Length 340;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6

Db 331 CEEDFY 336

RESULT 12

AAP92108

ID AAP92108 standard; protein: 637 AA.

XX AAP92108;

XX 14-FEB-1990 (first entry)

DT Human growth hormone receptor.

DE Growth hormone receptor.

XX Homo sapiens.

XX US4857637-A.

PN 15-AUG-1989.

XX 12-JUN-1987; 87US-0061942.

XX 07-MAY-1986; 86US-0861236.

PR 22-MAY-1985; 85US-0737302.

PR 12-JUN-1987; 87US-0061942.

XX (GETH) GENENTECH INC.

PA Hammonds RG, Leung DW, Martin DW, Spencer SA, Wood WI;

XX

XX

XX

XX

XX

XX

XX

DR WPI: 1989-300419/41.
DR N-PSDB: AAN91325.
XX
XX Modulating growth hormone receptor activity - by immunising animal
PT against growth hormone receptor extracellular domain deriv. to raise
PT antisera.
XX
XX
XX Disclosure; Fig. 2a-c; 18pp; English.
XX
XX An animal can be immunised against its growth hormone receptor by
CC vaccinating against a growth hormone receptor extracellular domain deriv.
CC predetermined to raise polyclonal antisera which affect the receptor as a
CC growth hormone agonist. This method enables continuous growth of target
CC tissues without frequent hormone admin.
XX
XX
SQ Sequence 637 AA;
Query Match 88.4%; Score 38; DB 10; Length 637;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEEDFY 6
Db 259 CEEDFY 264
RESULT 13
AAP81326
ID AAP81326 standard; protein; 638 AA.
XX
XX AAP81326;
XX
XX 23-OCT-1990 (first entry)
DT Human growth hormone receptor.
DE
XX Growth hormone receptor; gigantism; acromegaly.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..18
FT /label= signal_sequence
FT W08809818-A.
XX
XX 15-DEC-1988.
XX
XX 10-JUN-1988; 88WO-US02008.
XX
XX 12-JUN-1987; 87US-0062542.
XX
XX (GETH) GENENTECH INC.
XX
XX Hammonds RG, Leung DW, Spencer SA, Wood WI;
XX
XX WPI: 1988-368632/51.
DR N-PSDB: AAN81716.
XX
XX New pure growth hormone receptor and binding protein - for treating
PT growth hormone abnormalities, and new encoding DNA sequences.
XX
XX Disclosure; 1 pp; English.
XX
XX The sequence was deduced from a clone isolated from an adult liver
CC cDNA lambda gt10 library. The DNA can be inserted into an
CC expression vector for prodn. of the recombinant GHR which is used
CC to treat GH-related disorders such as gigantism and acromegaly.
CC A hydropathy plot revealed an extracellular GH binding domain, a
CC transmembrane domain, and an intracellular signalling domain. Eight
CC potential N-linked glycosylation sites are predicted.
CC See also AAP81327 and AAN81718-9.
XX

SQ Sequence 638 AA;
Query Match 88.4%; Score 38; DB 9; Length 638;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEEDFY 6
Db 259 CEEDFY 264
RESULT 14
AAW33394
ID AAW33394 standard; Protein: 638 AA.
XX
XX AC AAW33394;
XX
XX 11-MAY-1998 (first entry)
DT Human growth hormone receptor.
DE
XX Growth hormone receptor; growth hormone binding protein;
KW somatotropin; human; gigantism; acromegaly; therapy.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..18
FT /label= Sig_peptide
FT Domain 247..269
FT /note= "transmembrane domain"
FT Misc-difference 375
FT /note= "translated codon is Ser in clone ghr.210
FT and ghr.110, Ile in ghr.501"
XX
XX US5688763-A.
XX
XX 18-NOV-1997.
XX
XX 12-JUN-1987; 87US-0062542.
XX
XX 08-JAN-1993; 93US-0002489.
XX 12-JUN-1987; 87US-0062542.
XX 28-JUN-1991; 91US-0723358.
XX 25-MAY-1994; 94US-0248832.
XX
XX (COLO/) COLOSI P C.
XX (HAMM/) HAMMONDS R G.
XX (LEUNG/) LEUNG D W.
XX (SPEN/) SPENCER S A.
XX (WOOD/) WOOD W I.
XX
XX Colosi PC, Hammonds RG, Leung DW, Spencer SA, Wood WI;
XX
XX WPI: 1998-008010/01.
DR N-PSDB: AAT94063.
XX
XX Human and rabbit growth hormone receptor protein - useful to treat
PT disorders associated with overexpression, e.g. gigantism and
PT acromegaly
XX
XX Claim 2; Fig 8a; 60pp; English.
XX
XX This protein sequence comprises human growth hormone receptor. The
CC amino acid sequence was deduced from cDNA clones (see AAT94063)
CC obtained from a human liver cDNA library, and shows 84% identity
CC to the rabbit growth hormone receptor (see AAW33395). Human
CC growth hormone receptor, its derivatives in which the cytoplasmic
CC or transmembrane domains are deleted, and growth hormone binding
CC proteins comprising amino acids 190-246 or 1-324 of the mature
CC protein, can be used to treat disorders associated with growth
CC hormone over-expression, e.g. gigantism and acromegaly. The
CC binding protein may also be used to increase the stability and

CC efficacy of growth hormone in vivo.

XX Sequence 638 AA;

Query Match 88.4%; Score 38; DB 19; Length 638;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6

|||||

Db 259 CEEDFY 264

RESULT 15

ABBI1437

ID ABBI1437 standard; peptide; 648 AA.

XX AC ABBI1437;

DT 11-JAN-2002 (first entry)

XX Human growth hormone receptor homologue, SEQ ID NO:1807.

DE Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.

XX Homo sapiens.

XX WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

XX N-PSDB; ABA08681.

XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -

XX Claim 20; Page 189-190; 1963pp; English.

XX Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;

CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX SQ Sequence 648 AA;

Query Match 88.4%; Score 38; DB 22; Length 648;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6

|||||

Db 269 CEEDFY 274

Search completed: February 12, 2003, 11:44:39

Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 11:43:40 ; Search time 10.7333 Seconds
(without alignments)
19.189 Million cell updates/sec

Title: US-09-660-302C-7
Perfect score: 43
Sequence: 1 CEEDFYR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 38 | 88.4 | 620 | 4 | US-09-000-145-1 |
| 2 | 36 | 83.7 | 84 | 2 | US-08-465-380-59 |
| 3 | 36 | 83.7 | 84 | 2 | US-08-486-397-59 |
| 4 | 36 | 83.7 | 84 | 2 | US-08-486-399-59 |
| 5 | 36 | 83.7 | 84 | 2 | US-08-461-965-59 |
| 6 | 36 | 83.7 | 84 | 2 | US-08-634-641-59 |
| 7 | 36 | 83.7 | 84 | 3 | US-09-249-471-59 |
| 8 | 36 | 83.7 | 84 | 3 | US-09-249-472-59 |
| 9 | 36 | 83.7 | 84 | 3 | US-08-451-59 |
| 10 | 36 | 83.7 | 84 | 3 | US-08-809-455-59 |
| 11 | 36 | 83.7 | 84 | 3 | US-03-249-461-59 |
| 12 | 36 | 83.7 | 84 | 3 | US-09-249-448-59 |
| 13 | 36 | 83.7 | 91 | 2 | US-08-465-380-128 |
| 14 | 36 | 83.7 | 91 | 2 | US-08-480-478-50 |
| 15 | 36 | 83.7 | 91 | 2 | US-08-486-397-128 |
| 16 | 36 | 83.7 | 91 | 2 | US-08-486-399-128 |
| 17 | 36 | 83.7 | 91 | 2 | US-08-461-965-128 |
| 18 | 36 | 83.7 | 91 | 2 | US-08-326-110A-50 |
| 19 | 36 | 83.7 | 91 | 2 | US-08-634-641-128 |
| 20 | 36 | 83.7 | 91 | 3 | US-09-249-471-128 |
| 21 | 36 | 83.7 | 91 | 3 | US-09-249-472-128 |
| 22 | 36 | 83.7 | 91 | 3 | US-09-249-451-128 |
| 23 | 36 | 83.7 | 91 | 3 | US-08-809-455-128 |
| 24 | 36 | 83.7 | 91 | 3 | US-09-249-461-128 |
| 25 | 36 | 83.7 | 91 | 3 | US-09-249-448-128 |
| 26 | 34 | 79.1 | 723 | 6 | 5200183-4 |
| 27 | 33 | 76.7 | 75 | 2 | US-08-465-380-6 |

| | | | | | | |
|----|----|------|----|---|-------------------|-------------------|
| 28 | 33 | 76.7 | 75 | 2 | US-08-465-380-41 | Sequence 41, Appl |
| 29 | 33 | 76.7 | 75 | 2 | US-08-480-478-35 | Sequence 35, Appl |
| 30 | 33 | 76.7 | 75 | 2 | US-08-486-397-6 | Sequence 6, Appli |
| 31 | 33 | 76.7 | 75 | 2 | US-08-486-397-41 | Sequence 41, Appl |
| 32 | 33 | 76.7 | 75 | 2 | US-08-486-399-6 | Sequence 6, Appli |
| 33 | 33 | 76.7 | 75 | 2 | US-08-486-399-41 | Sequence 41, Appl |
| 34 | 33 | 76.7 | 75 | 2 | US-08-461-965-6 | Sequence 6, Appli |
| 35 | 33 | 76.7 | 75 | 2 | US-08-461-965-41 | Sequence 41, Appl |
| 36 | 33 | 76.7 | 75 | 2 | US-08-326-110A-35 | Sequence 35, Appl |
| 37 | 33 | 76.7 | 75 | 2 | US-08-634-641-6 | Sequence 6, Appli |
| 38 | 33 | 76.7 | 75 | 2 | US-08-634-641-41 | Sequence 41, Appl |
| 39 | 33 | 76.7 | 75 | 3 | US-09-249-471-6 | Sequence 6, Appli |
| 40 | 33 | 76.7 | 75 | 3 | US-09-249-471-41 | Sequence 41, Appl |
| 41 | 33 | 76.7 | 75 | 3 | US-09-249-472-6 | Sequence 6, Appli |
| 42 | 33 | 76.7 | 75 | 3 | US-09-249-472-41 | Sequence 41, Appl |
| 43 | 33 | 76.7 | 75 | 3 | US-09-249-451-6 | Sequence 6, Appli |
| 44 | 33 | 76.7 | 75 | 3 | US-09-249-451-41 | Sequence 41, Appl |
| 45 | 33 | 76.7 | 75 | 3 | US-08-809-455-6 | Sequence 6, Appli |

ALIGNMENTS

RESULT 1

US-09-000-145-1
; Sequence 1, Application US/09000145
; Patent No. 6169172
; GENERAL INFORMATION:
; APPLICANT: DEVAUCHELLE, Gerrard
; APPLICANT: GARNIER, Laurence
; APPLICANT: CAHOREAU, Claire
; APPLICANT: CERUTTI, Martine
; TITLE OF INVENTION: USE OF A PROLACTIN RECEPTOR OR GROWTH HORMONE RECEPTOR
; FILE REFERENCE: INTRACYTOPLASMIC DOMAIN FOR ACHIEVING PROTEIN SECRETION
; CURRENT APPLICATION NUMBER: US/09/000.145
; CURRENT FILING DATE: 1998-03-16
; EARLIER APPLICATION NUMBER: PCT/FR96/01237
; EARLIER FILING DATE: 1996-08-02
; EARLIER APPLICATION NUMBER: FR 95/09420
; EARLIER FILING DATE: 1995-08-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-000-145-1

Query Match 88.4%; Score 38; DB 4; Length 620;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative; 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6
| | | | |
Db 241 CEEDFY 246

RESULT 2

US-08-465-380-59
; Sequence 59, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Menssens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,380
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-465-380-59

Query Match 83.7%; Score 36; DB 2; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 CEEDFYR 7
||| |||
Db 52 CEEGFYR 58

RESULT 3
US-08-486-397-59
Sequence 59, Application US/08486397
Patent No. 5866542
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 357
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59

Query Match 83.7%; Score 36; DB 2; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 CEEDFYR 7
||| |||
Db 52 CEEGFYR 58

RESULT 4
US-08-486-399-59
Sequence 59, Application US/08486399
Patent No. 5866543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,399
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/270
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-399-59

Query Match 83.7%; Score 36; DB 2; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEEGFYR 7
Db 52 CEEGFYR 58

RESULT 5

US-08-461-965-59
Sequence 59, Application US/08461965
Patent No. 5872098
GENERAL INFORMATION:

APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,965
FILING DATE: June 5, 1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 210/243
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-461-965-59

Query Match 83.7%; Score 36; DB 2; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEEGFYR 7
Db 52 CEEGFYR 58

RESULT 6

US-08-634-641-59
Sequence 59, Application US/08634641
Patent No. 5955294
GENERAL INFORMATION:

APPLICANT: Vlasuk, George P. Vlasuk
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Mensens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/634,641
FILING DATE: April 19, 1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 219/136
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum

US-08-634-641-59

Query Match 83.7%; Score 36; DB 2; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEEDFYR 7

Db 52 CEEGFYR 58

RESULT 7

US-09-249-471-59
; Sequence 59, Application US/09249471
; Patent No. 604041

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Suite 4700

STATE: Los Angeles

COUNTRY: California

ZIP: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/249,471

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/809,455

FILING DATE: April 17, 1997

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486,399

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/486,397

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/461,965

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 216/270

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: Ancylostoma caninum

US-09-249-471-59

Query Match

Best Local Similarity 83.7%; Score 36; DB 3; Length 84;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEEDFYR 7

Db 52 CEEGFYR 58

RESULT 8

US-09-249-472-59
; Sequence 59, Application US/09249472
; Patent No. 6046318

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE

TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Suite 4700

STATE: Los Angeles

COUNTRY: California

ZIP: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/249,472

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/809,455

FILING DATE: April 17, 1997

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486,399

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/486,397

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/461,965

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 216/270

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-09-249-472-59

Query Match 83.7%; Score 36; DB 3; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEEGFYR 7

|||||

Db 52 CEEGFYR 58

RESULT 9

US-09-249-451-59
Sequence 59, Application US/09249451
Patent No. 6087487

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansmans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,451
FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270

TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-09-249-451-59

Query Match 83.7%; Score 36; DB 3; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEEGFYR 7

|||||

Db 52 CEEGFYR 58

RESULT 10

US-08-809-455-59
Sequence 59, Application US/08809455
Patent No. 6090916

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansmans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,455
FILING DATE: April 17, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.

; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-08-809-455-59

Query Match 83.7%; Score 36; DB 3; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 CEEGFYR 7
||| |||
Db 52 CEEGFYR 58

RESULT 11

US-09-249-461-59
; Sequence 59, Application US/09249461
; Patent No. 6096877

GENERAL INFORMATION:

; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Larocher, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,461
; FILING DATE:

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995

; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-09-249-461-59

Query Match 83.7%; Score 36; DB 3; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 CEEGFYR 7
||| |||
Db 52 CEEGFYR 58

RESULT 12

US-09-249-448-59
; Sequence 59, Application US/09249448
; Patent No. 6121435

GENERAL INFORMATION:

; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Larocher, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,448
; FILING DATE:
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-09-249-448-59

Query Match 83.7%; Score 36; DB 3; Length 84;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CREDFYR 7
Db 52 CEEGFYR 58

RESULT 13
US-08-465-380-128
Sequence 128, Application US/08465380
Patent No. 5863894
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H. L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,380
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 128:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-465-380-128

Query Match 83.7%; Score 36; DB 2; Length 91;
Best Local Similarity 85.7%; Pred. No. 5.4;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CREDFYR 7
Db 59 CEEGFYR 65

RESULT 14
US-08-480-478-50
Sequence 50, Application US/08480478
Patent No. 5864009
GENERAL INFORMATION:
APPLICANT: GEORGE P. VLASUK; PATRICK ERIC
APPLICANT: HUGO STANSSENS; JORIS HILDA
APPLICANT: LIEVEN MESSENS; MARC JOZEF
APPLICANT: LAUWEREYS; YVES RENE LAROCHE;
APPLICANT: LAURENT STEPHANE JESPEERS; and
APPLICANT: YANNICK GEORGES JOZEF
APPLICANT: GANSEMANS
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-
TITLE OF INVENTION: COAGULANT PROTEIN
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,478
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: 18 OCTOBER 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 208/290
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-480-478-50

Query Match 83.7%; Score 36; DB 2; Length 91;

Best Local Similarity 85.7%; Pred. No. 5.4;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEEDFYR 7
 | | | | |
Db 59 CEEGFYR 65

RESULT 15
US-08-486-397-128
; Sequence 128: Application US/08486397
; Patent No. 5866542
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 357
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon 6 Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-08-486-397-128

Query Match 83.7%; Score 36; DB 2; Length 91;
Best Local Similarity 85.7%; Pred. No. 5.4;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEEDFYR 7
 | | | | |
Db 59 CEEGFYR 65

Search completed: February 12, 2003, 11:48:01
Job time : 11.7333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 11:39:59 ; Search time 6.53333 Seconds
(without alignments)
44.439 Million cell updates/sec

Title: US-09-660-302C-7
Perfect score: 43
Sequence: 1 CEEDFYR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|--------------|---------------------|
| 1 | 38 | 88.4 | 638 | 1 GHR_HUMAN | P10912 homo sapien |
| 2 | 38 | 88.4 | 638 | 1 GHR_MACMU | P79194 macaca mula |
| 3 | 36 | 83.7 | 20 | 1 UN05_PINPS | P81674 pinus pinas |
| 4 | 36 | 83.7 | 364 | 1 MYCL_HUMAN | P12524 homo sapien |
| 5 | 36 | 83.7 | 368 | 1 MYCL_MOUSE | P10166 mus musculus |
| 6 | 36 | 83.7 | 388 | 1 LHX9_HUMAN | O9nq69 homo sapien |
| 7 | 36 | 83.7 | 388 | 1 LHX9_MOUSE | O9wh2 mus musculus |
| 8 | 36 | 83.7 | 406 | 1 LHX2_HUMAN | P50458 homo sapien |
| 9 | 36 | 83.7 | 406 | 1 LHX2_MOUSE | O920s2 mus musculus |
| 10 | 36 | 83.7 | 426 | 1 LHX2_RAT | P36198 rattus norv |
| 11 | 35 | 81.4 | 372 | 1 COLB_ARATH | O9sse5 arabidopsis |
| 12 | 34 | 79.1 | 599 | 1 BAL_MOUSE | O64285 mus musculus |
| 13 | 34 | 79.1 | 612 | 1 BAL_RAT | P07882 rattus norv |
| 14 | 34 | 79.1 | 707 | 1 NRDD_HAETN | P43752 haemophilus |
| 15 | 33 | 76.7 | 86 | 1 Y576_ARCFU | O29679 archaeoglob |
| 16 | 32 | 74.4 | 154 | 1 PYR1_BUCAI | P57451 buchnera ap |
| 17 | 32 | 74.4 | 367 | 1 LHX4_HUMAN | O969q2 homo sapien |
| 18 | 32 | 74.4 | 367 | 1 LHX4_MOUSE | P53776 mus musculus |
| 19 | 32 | 74.4 | 395 | 1 LHX3_CHICK | P53412 gallus gall |
| 20 | 32 | 74.4 | 415 | 1 THD1_BACHD | O9kc63 bacillus ha |
| 21 | 32 | 74.4 | 486 | 1 HH1R_RAT | P31390 rattus norv |
| 22 | 32 | 74.4 | 487 | 1 HH1R_HUMAN | P35367 homo sapien |
| 23 | 32 | 74.4 | 488 | 1 HH1R_CAVPO | P31389 cavia porce |
| 24 | 32 | 74.4 | 488 | 1 HH1R_MOUSE | P70174 mus musculus |
| 25 | 32 | 74.4 | 491 | 1 HH1R_BOVIN | P30546 bos taurus |
| 26 | 32 | 74.4 | 581 | 1 NET2_CHICK | O90923 gallus gall |
| 27 | 32 | 74.4 | 606 | 1 NET1_CHICK | O90922 gallus gall |
| 28 | 32 | 74.4 | 647 | 1 PPPO_LACHE | O52071 lactobacill |
| 29 | 32 | 74.4 | 647 | 1 SKO1_YEAST | O02100 saccharomyc |
| 30 | 32 | 74.4 | 727 | 1 NET4_DROME | O24567 drosophila |
| 31 | 32 | 74.4 | 796 | 1 COPP_SCHPO | O42937 schizosacch |
| 32 | 32 | 74.4 | 976 | 1 EPA2_HUMAN | P29317 homo sapien |
| 33 | 32 | 74.4 | 1005 | 1 EPA5_RAT | P54757 rattus norv |

| | | | | | |
|----|----|------|------|--------------|--------------------|
| 34 | 32 | 74.4 | 1037 | 1 EPA5_HUMAN | P54756 homo sapien |
| 35 | 32 | 74.4 | 1696 | 1 PKC5_BRACL | O9nj15 branchiosto |
| 36 | 31 | 72.1 | 90 | 1 YHCO_ECOLI | P46480 escherichia |
| 37 | 31 | 72.1 | 148 | 1 FER2_ARATH | O04090 arabidopsis |
| 38 | 31 | 72.1 | 230 | 1 YGEA_ECOLI | P03813 escherichia |
| 39 | 31 | 72.1 | 334 | 1 Y10A_BPT4 | P32277 bacterioph |
| 40 | 31 | 72.1 | 385 | 1 YHH7_SACDO | P41907 saccharomyc |
| 41 | 31 | 72.1 | 401 | 1 T11B_HUMAN | O00300 homo sapien |
| 42 | 31 | 72.1 | 401 | 1 T11B_MOUSE | O08712 mus musculu |
| 43 | 31 | 72.1 | 401 | 1 T11B_RAT | O08727 rattus norv |
| 44 | 31 | 72.1 | 597 | 1 BAL_BOVIN | P30122 bos taurus |
| 45 | 31 | 72.1 | 621 | 1 GIDA_HELPY | P56138 helicobacte |

ALIGNMENTS

RESULT 1
GHR_HUMAN STANDARD; PRT; 638 AA.
AC P10912;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Growth hormone receptor precursor (GH receptor) (Serum binding
DE protein).
GN GHR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=88065896; PubMed=2825030;
RA Leung D.W., Spencer S.A., Cachianes G., Hammonds R.G., Collins C.,
RA Henzel W.J., Barnard R., Waters M.J., Wood W.I.;
RT "Growth hormone receptor and serum binding protein: purification,
RT cloning and expression";
RL Nature 330:537-543(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90046742; PubMed=2813379;
RA Godowski P.J., Leung D.W., Meacham L.R., Galgani J.P., Hellmiss R.,
RA Keret R., Rotwein P.S., Parks J.S., Laron Z., Wood W.I.;
RT "Characterization of the human growth hormone receptor gene and
RT demonstration of a partial gene deletion in two patients with Laron-
RT type dwarfism.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8083-8087(1989).
RN [3]
RP DISULFIDE BONDS.
RX MEDLINE=90153957; PubMed=2406245;
RA Fuh G., Mulikerrin M.G., Bass S., McFarland N., Brochier M.,
RA Bourrel J.H., Light D.R., Wells J.A.;
RT "The human growth hormone receptor. Secretion from Escherichia coli
RT and disulfide bonding pattern of the extracellular binding domain.";
RL J. Biol. Chem. 265:3111-3115(1990).
RN [4]
RP VARIANT LARON DWARFISM SER-114.
RX MEDLINE=89384829; PubMed=2779634;
RA Amselem S., Duquesnoy P., Attree O., Novelli G., Bousnina S.,
RA Postelvinay M.-C., Goossens M.;
RT "Laron dwarfism and mutations of the growth hormone-receptor gene.";
RL New Engl. J. Med. 321:989-995(1989).
RN [5]
RP VARIANTS LARON DWARFISM.
RX MEDLINE=93278381; PubMed=8504296;
RA Amselem S., Duquesnoy P., Duriez B., Dastot F., Sorbier M.-L.,
RA Valleix S., Goossens M.;
RT "Spectrum of growth hormone receptor mutations and associated
RT haplotypes in Laron syndrome.";
RL Hum. Mol. Genet. 2:355-359(1993).
RN [6]

RP VARIANT LARON DWARFISM HIS-170.
RX MEDLINE=94185645; PubMed=9137822;
RA Duesquesnoy P., Sobrier M.-L., Dastot F., Buchanan C.R.,
RA Savage M.O., Preece M.A., Craescu C.T., Blouquit Y., Goossens M.,
RA Amselme S.;
RT "A single amino acid substitution in the exoplasmic domain of the
RT human growth hormone (GH) receptor confers familial GH resistance
RT (Laron syndrome) with positive GH-binding activity by abolishing
RT receptor homodimerization.";
RL EMBO J. 13:1386-1395(1994).
RN [7]
RN VARIANTS IDIOPATHIC SHORT STATURE LYS-62; CYS-179 AND ASP-242.
RX MEDLINE=96013502; PubMed=7565946;
RA Goddard A.D., Covello R., Luch S.-M., Clarkson T., Attie K.M.,
RA Gundhardt N., Rundle A.C., Wells J.A., Carlsson L.M.S.;
RT "Mutations of the growth hormone receptor in children with idiopathic
RT short stature.";
RL New Engl. J. Med. 333:1093-1098(1995).
RN [8]
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 19-256.
RX MEDLINE=92196577; PubMed=1549776;
RA de Vos A.M., Ullsch M., Kossiakoff A.A.;
RT "Human growth hormone and extracellular domain of its receptor:
RT crystal structure of the complex.";
RL Science 255:306-312(1992).
RN [9]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 19-256.
RX MEDLINE=97113023; PubMed=8943276;
RA Sundstroem M., Lundqvist T., Roedin J., Giebel L.B., Milligan D.,
RA Norstedt G.;
RT "Crystal structure of an antagonist mutant of human growth hormone,
RT G120R, in complex with its receptor at 2.9-A resolution.";
RL J. Biol. Chem. 271:32197-32203(1996).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DISEASE: DEFICIENCY IN GHR IS THE CAUSE OF PITUITARY DWARFISM II
CC (LARON-TYPE PITUITARY DWARFISM OR LARON SYNDROME (LS)). IT ALSO
CC CAUSES IDIOPATHIC SHORT STATURE.
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE TYPE III-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06562; CAA29808.1; -
CC EMBL; M28466; AAA52555.1; -
CC EMBL; M28458; AAA52555.1; JOINED.
CC EMBL; M28459; AAA52555.1; JOINED.
CC EMBL; M28460; AAA52555.1; JOINED.
CC EMBL; M28461; AAA52555.1; JOINED.
CC EMBL; M28462; AAA52555.1; JOINED.
CC EMBL; M28463; AAA52555.1; JOINED.
CC EMBL; M28464; AAA52555.1; JOINED.
CC EMBL; M28465; AAA52555.1; JOINED.
CC PIR; S04530; S04530.
CC PIR; A33991; A33991.
CC PDB; 3HHR; 30-APR-94.
CC PDB; 1HWG; 19-NOV-97.
CC PDB; 1WHH; 19-NOV-97.
CC PDB; 1AXI; 28-JAN-98.
CC PDB; 1A22; 29-APR-98.
CC Genew; HGNC:4263; GHR.
CC MIM; 600946; -
CC MIM; 262500; -
CC InterPro; IPR002996; CRLA.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR003528; Hemtopopn_L_F1.

DR Pfam; PF00041; fn3; 1.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal; 3D-structure;
KW Dwarfism; Disease mutation.
FT SIGNAL 1 18
FT CHAIN 19 638
FT DOMAIN 19 264
FT TRANSMEM 265 288
FT DOMAIN 289 638
FT DOMAIN 145 252
FT DISULFID 56 66
FT DISULFID 101 112
FT DISULFID 126 140
FT CARBOHYD 46 46
FT CARBOHYD 115 115
FT CARBOHYD 156 156
FT CARBOHYD 161 161
FT CARBOHYD 200 200
FT VARIANT 62 62
FT VARIANT 89 89
FT VARIANT 114 114
FT VARIANT 143 143
FT VARIANT 162 162
FT VARIANT 170 170
FT VARIANT 179 179
FT VARIANT 229 229
FT VARIANT 242 242
FT CONFLICT 544 544
FT STRAND 53 58
FT STRAND 64 68
FT STRAND 82 88
FT STRAND 99 100
FT TURN 104 107
FT TURN 109 110
FT STRAND 111 114
FT TURN 116 117
FT STRAND 124 131
FT TURN 132 133
FT STRAND 134 142
FT HELIX 143 145
FT STRAND 147 147
FT STRAND 153 162
FT TURN 164 165
FT STRAND 168 176
FT TURN 179 180
FT TURN 183 186
FT STRAND 190 198
FT TURN 199 200
FT STRAND 205 206
FT STRAND 210 210
FT STRAND 214 221
FT TURN 222 223
FT STRAND 225 234
FT STRAND 247 250
SQ SEQUENCE 638 AA; 71499 MW; EAF77EAD64787822 CRC64;
Query Match 88.4%; Score 38; DB 1; Length 638;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEDFY 6

```
Db 259 CEEDFY 264
|||||
RESULT 2
GHR_MACMU
ID GHR_MACMU STANDARD; PRT; 638 AA.
AC P79194;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Growth hormone receptor precursor (GH receptor) (Serum binding
protein).
DE GHR.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97373601; PubMed=9228076;
RA Martini J.F., Pezet A., Guezennec C.Y., Edery M., Postel-Vinay M.C.,
RA Kelly P.A.;
RT "Monkey growth hormone (GH) receptor gene expression. Evidence for
RT two mechanisms for the generation of the GH binding protein.";
RL J. Biol. Chem. 272:18951-18958(1997).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U84589; AAB47702.1; -
DR HSSP; P10912; IAXI.
DR InterPro; IPR002996; CRIA.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR003528; Hemtopoptn_L_F1.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 638 GROWTH HORMONE RECEPTOR.
FT DOMAIN 19 264 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 265 288 POTENTIAL.
FT DOMAIN 289 638 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 145 252 FIBRONECTIN TYPE-III.
FT DISULFID 56 66 BY SIMILARITY.
FT DISULFID 101 112 BY SIMILARITY.
FT DISULFID 126 140 BY SIMILARITY.
SQ SEQUENCE 638 AA; 71327 MW; 1F81A55301625F8E CRC64;

Query Match 88.4%; Score 38; DB 1; Length 638;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6
|||||
Db 259 CEEDFY 264

RESULT 3
UN05_PINPS
ID UN05_PINPS STANDARD; PRT; 20 AA.
AC P81674;
```

```
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2b-page of needles (N147) (Fragments).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RT TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.6, ITS MW IS: 36 kDa.
FT NON_TER 1 1
FT NON_CONS 11 12 F -> I.
FT VARIANT 13 13 Y -> E.
FT VARIANT 14 14 R -> K.
FT VARIANT 15 15
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2438 MW; 9F4B4678E086C298 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 20;
Best Local Similarity 71.4%; Pred. No. 0.43;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 7
|||||
Db 9 CDKDFYR 15
|||||

RESULT 4
MYCL_HUMAN
ID MYCL_HUMAN STANDARD; PRT; 364 AA.
AC P12524; Q9NUE9;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE L-myc-1 proto-oncogene protein.
GN MYCL1 OR MYCL OR LMVC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88094386; PubMed=2827002;
RA Kaye F., Battey J., Nau M., Brooks B., Seifter E., de Greve J.,
RA Birrer M., Sausville E., Minna J.;
RT "Structure and expression of the human L-myc gene reveal a complex
RT pattern of alternative mRNA processing.";
RL Mol. Cell. Biol. 8:186-195(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89112807; PubMed=3322939;
RA DePinho R.A., Hatton K.S., Testfaye A., Yancopoulos G.D., Alt F.W.;
RT "The human myc gene family: structure and activity of L-myc and an
RT L-myc pseudogene.";
RL Genes Dev. 1:1311-1326(1987).
RN [3]
RP SEQUENCE FROM N.A.
RA Ellington A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M19720; AAA59879.1; -;
 DR EMBL; X07262; CAA30248.1; -;
 DR EMBL; X07263; CAA30249.1; -;
 DR EMBL; AL033527; CAB75682.1; -;
 DR PIR; A27675; TVHUML.
 DR HSSP; P25912; IHLO.
 DR TRANSFAC; T02385; -;
 DR Genew; HGNC:7555; MYCL1.
 DR MIW; 164850; -;
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR002418; TF_Myc.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF01056; Myc_N_term; 1.
 DR PRINTS; PR00044; LEUZIPPRMYC.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS00888; HLH_2; 1.
 KW Nuclear protein; DNA-binding; Proto-oncogene.
 FT DNA_BIND 282 294
 FT DOMAIN 295 334
 FT DOMAIN 333 361
 FT CONFLICT 362 362 S -> T (IN REF. 3).
 SQ SEQUENCE 364 AA; 40312 MW; 58F8A71A1C2ED6D4 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 364;
 Best Local Similarity 85.7%; Pred. No. 8.3;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
 I I I I I
 Db 15 CGEDFYR 21

RESULT 5
 MYCL_MOUSE STANDARD; PRT; 368 AA.
 ID MYCL_MOUSE
 AC P10166;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-myc proto-oncogene protein.
 GN MYCL1 OR MYCL OR LMYC1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/c;
 RX MEDLINE=88111523; PubMed=2828024;
 RA Legoux E., Depinho R.A., Zimmerman K., Collum R., Yancopoulos G.D.,
 RA Mitsock L., Kriz R., Alt F.W.;
 RT "Structure and expression of the murine L-myc gene.";
 RL EMBO J. 6:3359-3366(1987).
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X13945; CAA32128.1; -;
 CC PIR; S03017; TVMSML.
 CC HSSP; P25912; IHLO.
 CC TRANSFAC; T02387; -;
 CC MGI; 96799; Lmycl1.
 CC InterPro; IPR001092; HLH_basic.
 CC InterPro; IPR002418; TF_Myc.
 CC Pfam; PF00010; HLH; 1.
 CC Pfam; PF01056; Myc_N_term; 1.
 CC PRINTS; PR00044; LEUZIPPRMYC.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00038; HLH_1; 1.
 CC PROSITE; PS00888; HLH_2; 1.
 KW Nuclear protein; DNA-binding; Proto-oncogene.
 FT DNA_BIND 286 298
 FT DOMAIN 299 338
 FT DOMAIN 337 365
 SQ SEQUENCE 368 AA; 40848 MW; 9174F6FD7C03321E CRC64;

Query Match 83.7%; Score 36; DB 1; Length 368;
 Best Local Similarity 85.7%; Pred. No. 8.3;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
 I I I I I
 Db 15 CGEDFYR 21

RESULT 6
 LHX9_HUMAN STANDARD; PRT; 388 AA.
 ID LHX9_HUMAN
 AC Q9NQ69; Q9NQ70; Q9BYU6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE LIM/homeobox protein Lhx9.
 GN LHX9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=21291005; PubMed=11397841;
 RA Ottolenghi C., Moreira-Filho C., Mendonca B.B., Barbieri M.,
 RA Fellous M., Berkovitz G.D., McElreavey K.;
 RT "Absence of mutations involving the LIM homeobox domain gene LHX9 in
 RT 46,XY gonadal agenesis and dysgenesis.";
 RL J. Clin. Endocrinol. Metab. 86:2465-2469(2001).
 CC -1- FUNCTION: INVOLVED IN GONADAL DEVELOPMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AJ277915; CAB97493.1; -;
 DR EMBL; AJ277916; CAB98128.1; ALT SEQ.
 DR EMBL; AJ277917; CAB98128.1; JOINED.
 DR EMBL; AJ277918; CAB98128.1; JOINED.
 DR EMBL; AJ277919; CAB98128.1; JOINED.
 DR EMBL; AJ277920; CAB98128.1; JOINED.
 DR EMBL; AJ296272; CAC33174.1; -;


```
DR HSP: P06601; 1FJL.
DR Genew; HGNC:14222; LHX9.
DR MIM; 606066; -.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF04112; LIM; 2.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000094; LIM; 2.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS00023; LIM_DOMAIN_2; 2.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein; Repeat; LIM domain;
KW Metal-binding; Zinc.
KT DOMAIN 62 114 LIM 1.
FT DOMAIN 124 177 LIM 2.
FT DNA_BIND 258 317 HOMEBOX.
SQ SEQUENCE 388 AA; 42903 MW; A4DC8B914D7C3B66 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 388;
Best Local Similarity 71.4%; Pred. No. 8.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
Db 111 CKEDYIR 117

RESULT 7
LHX9_MOUSE
ID LHX9_MOUSE STANDARD; PRT; 388 AA.
AC Q9WUH2; Q9WUH4; Q9QYQ5; Q9QYQ6; Q9QZ00;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE LIM/homeobox protein Lhx9.
GN LHX9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-300 FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=9908964; PubMed=9880598;
RA Retazzi S., Rogard M., Bach I., Failli V., Besson M.J.;
RT "Lhx9: a novel LIM-homeodomain gene expressed in the developing
RT forebrain.";
RL J. Neurosci. 19:783-793(1999).
RN [2]
RP SEQUENCE OF 11-388 FROM N.A.
RC STRAIN=NIH Swiss;
RX MEDLINE=99264291; PubMed=10330499;
RA Bertuzzi S., Porter F.D., Pitts A., Kumar M., Agulnick A., Wassif C.,
RA Westphal H.;
RT "Characterization of Lhx9, a novel LIM/homeobox gene expressed by the
RT pioneer neurons in the mouse cerebral cortex.";
RL Mech. Dev. 81:193-198(1999).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=20221375; PubMed=10756098;
RA Failli V., Rogard M., Mattei M.-G., Vernier P., Retaux S.;
RT "Lhx9 and Lhx9alpha LIM-homeodomain factors: genomic structure,
RT expression pattern, chromosomal localization, and phylogenetic
RT analysis.";
RL Genomics 64:307-317(2000).
CC -!- FUNCTION: INVOLVED IN GONADAL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA AND BETA (SHOWN HERE); ARE
```

```
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF134761; AAD30110.1; -
DR EMBL; AF113518; AAD22008.1; -
DR EMBL; AJ243851; CAB59907.1; -
DR EMBL; AJ243852; CAB59908.1; -
DR EMBL; AJ243853; CAB59908.1; JOINED.
DR EMBL; AJ243854; CAB59908.1; JOINED.
DR EMBL; AJ243855; CAB59908.1; JOINED.
DR EMBL; AJ243856; CAB59908.1; JOINED.
DR EMBL; AJ243852; CAB59909.1; -
DR EMBL; AJ243853; CAB59909.1; JOINED.
DR EMBL; AJ243854; CAB59909.1; JOINED.
DR EMBL; AJ243855; CAB59909.1; JOINED.
DR EMBL; AJ243857; CAB59909.1; JOINED.
DR HSP; P06601; 1FJL.
DR TRANSFAC; T04192; -.
DR TRANSFAC; T04195; -.
DR MGI; MGI:1316721; Lhx9.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF04112; LIM; 2.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000094; LIM; 2.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS00023; LIM_DOMAIN_2; 2.
KW Homeobox; DNA-binding; Nuclear protein; Repeat; LIM domain;
KW Metal-binding; Zinc; Alternative splicing.
FT DOMAIN 62 114 LIM 1.
FT DOMAIN 124 177 LIM 2.
FT DNA_BIND 258 317 HOMEBOX.
FT VARSPLIC 304 388
FT
FT CONFLICT 49 49 A -> T (IN REF. 2).
FT CONFLICT 153 153 S -> F (IN REF. 2).
SQ SEQUENCE 388 AA; 42986 MW; C2D7326A68D87B32 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 388;
Best Local Similarity 71.4%; Pred. No. 8.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
Db 111 CKEDYIR 117

RESULT 8
LHX2_HUMAN
ID LHX2_HUMAN STANDARD; PRT; 406 AA.
AC P50458; O55860;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE LIM/homeobox protein Lhx2 (Homeobox protein Lh2-2).
GN LHX2 OR LH2.
```

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96226351; PubMed=8649822;
 RA Wu H.-K., Heng H.H.O., Siderovski D.P., Dong W.-F., Okuno Y.,
 RA Shi X.-M., Tsui L.-C., Minden M.D.;
 RT "Identification of a human LIM-Hox gene, hLH-2, aberrantly expressed
 RT in chronic myelogenous leukaemia and located on 9q33-34.1.";
 RL Oncogene 12:1205-1212(1996).
 RN [2]
 RP SEQUENCE OF 18-406 FROM N.A.
 RX MEDLINE=99162575; PubMed=10051612;
 RA Rincon-Limas D.E., Lu C.-H., Canal I., Calleja M.,
 RA Rodriguez-Esteban C., Izpisua-Belmonte J.C., Botas J.;
 RT "Conservation of the expression and function of apterous orthologs in
 RT Drosophila and mammals.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2165-2170(1999).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATION INVOLVED IN THE
 CC CONTROL OF CELL DIFFERENTIATION IN DEVELOPING LYMPHOID AND
 CC NEURAL CELL TYPES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U11701; AAB08752.1; .
 CC EMBL; AF124735; AAD20013.1; .
 CC HSP; P06601; 1FJL.
 CC Genew; HGNC:6594; LHX2.
 CC MIM; 603759; .
 CC InterPro; IPR001356; Homeobox.
 CC InterPro; IPR001781; LIM.
 CC Pfam; PF00046; homeobox; 1.
 CC Pfam; PF00412; LIM; 2.
 CC ProDom; PD000010; Homeobox; 1.
 CC ProDom; PD000094; LIM; 2.
 CC SMART; SM00389; HOX; 1.
 CC SMART; SM00132; LIM; 2.
 CC PROSITE; PS00478; LIM_DOMAIN_1; 2.
 CC PROSITE; PS00023; LIM_DOMAIN_2; 2.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC Homeobox; DNA-binding; Nuclear protein; Repeat; LIM domain;
 CC Metal-binding; Zinc; Transcription regulation.
 CC DOMAIN 53 105
 CC LIM 1.
 CC FT DOMAIN 115 168
 CC POLY-ALA.
 CC FT DOMAIN 187 196
 CC HOMEBOX.
 CC FT DOMAIN 266 325
 CC NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DOMAIN 307 323
 CC AKSEAPA -> QERGR (IN REF. 1).
 CC A -> G (IN REF. 1).
 CC A -> Q (IN REF. 1).
 CC AAA -> RAR (IN REF. 1).
 CC MISSING (IN REF. 1).
 CC NA -> TR (IN REF. 1).
 CC GHFPHSPSTLTNLF -> AMSLTAPHKRLTLTFNSDSOP
 CC PHTISLKKLSLV (IN REF. 1).
 CC FT CONFLICT 391 406
 CC SEQUENCE 406 AA; 44373 MW; DF73AAC81867D30D CRC64;
 Query Match 83.7%; Score 36; DB 1; Length 406;
 Best Local Similarity 71.4%; Pred. No. 9.2;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEEDFYR 7
 Db 102 CKEDYR 108
 RESULT 9
 LHX2_MOUSE
 ID LHX2_MOUSE STANDARD; PRT; 406 AA.
 AC Q9Z0S2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE LIM/homeobox protein Lhx2.
 GN LHX2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99162575; PubMed=10051612;
 RA Rincon-Limas D.E., Lu C.-H., Canal I., Calleja M.,
 RA Rodriguez-Esteban C., Izpisua-Belmonte J.C., Botas J.;
 RT "Conservation of the expression and function of apterous orthologs in
 RT Drosophila and mammals.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2165-2170(1999).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATION INVOLVED IN THE
 CC CONTROL OF CELL DIFFERENTIATION IN DEVELOPING LYMPHOID AND
 CC NEURAL CELL TYPES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF124734; AAD20012.1; .
 CC HSP; P06601; 1FJL.
 CC TRANSFAC; T01969; .
 CC MGD; MGI:96785; Lhx2.
 CC InterPro; IPR001356; Homeobox.
 CC InterPro; IPR001781; LIM.
 CC Pfam; PF00046; homeobox; 1.
 CC Pfam; PF00412; LIM; 2.
 CC ProDom; PD000010; Homeobox; 1.
 CC ProDom; PD000094; LIM; 2.
 CC SMART; SM00389; HOX; 1.
 CC SMART; SM00132; LIM; 2.
 CC PROSITE; PS00478; LIM_DOMAIN_1; 2.
 CC PROSITE; PS00023; LIM_DOMAIN_2; 2.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC Homeobox; DNA-binding; Nuclear protein; Repeat; LIM domain;
 CC Metal-binding; Zinc; Transcription regulation.
 CC DOMAIN 53 105
 CC LIM 1.
 CC FT DOMAIN 115 168
 CC POLY-ALA.
 CC FT DOMAIN 187 196
 CC HOMEBOX.
 CC FT DOMAIN 266 325
 CC NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DOMAIN 307 323
 CC FE7B4E76454D6A90 CRC64;
 CC SEQUENCE 406 AA; 44419 MW;
 Query Match 83.7%; Score 36; DB 1; Length 406;
 Best Local Similarity 71.4%; Pred. No. 9.2;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CEEDFYR 7
Db 102 CKEDYR 108

RESULT 10
LHX2_RAT
ID LHX2_RAT STANDARD; PRT; 426 AA.
AC P36198;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE LIM/homeobox protein Lhx2 (Homeobox protein LH-2).
GN LHX2 OR LHX2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
TX TISSUE=Brain;
RX MEDLINE=93126348; PubMed=7678338;
RA Xu Y., Baldassare M., Fisher P., Rathbun G., Oltz E.M.,
RA Yancopoulos G.D., Jessell T.M., Alt F.W.;
RT "LH-2: a LIM/homeobox domain gene expressed in developing lymphocytes and
RT neural cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:227-231(1993).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATORY PROTEIN INVOLVED IN THE
CC CONTROL OF CELL DIFFERENTIATION IN DEVELOPING LYMPHOID AND
CC NEURAL CELL TYPES.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: FOUND IN DISCRETE REGIONS OF THE DEVELOPING
CC CNS. PRIMARILY IN DIENCEPHALIC AND TELENCEPHALIC STRUCTURES AND
CC A SUBSET OF LYMPHOID TISSUES. ALSO FOUND IN EMBRYONIC SPINAL CHORD
CC AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN DEVELOPING LYMPHOCYTES AND
CC NEURAL CELLS. MAXIMAL EXPRESSION IS FOUND IN PRE-B LYMPHOCYTES.
CC -!- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L06804; -; NOT_ANNOTATED_CDS.
CC HSSP; P06601; 1FJL.
CC TRANSFAC; T01966; -.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR001781; LIM.
CC Pfam; PF00046; homeobox; 1.
CC Pfam; PF00412; LIM; 2.
CC ProDom; PD000010; Homeobox; 1.
CC ProDom; PD000094; LIM; 2.
CC SMART; SM00389; HOX; 1.
CC SMART; SM00132; LIM; 2.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00478; LIM_DOMAIN_1; 2.
CC PROSITE; PS50023; LIM_DOMAIN_2; 2.
CC PROSITE; PS50071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Nuclear protein; Repeat; LIM domain;
CC Metal-binding; Zinc; Transcription regulation.
CC DOMAIN 52 104
CC FT 114 167 LIM 1.
CC FT 114 167 HOMEBOX.
CC FT DNA_BIND 264 323
CC FT DOMAIN 305 321 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC SEQUENCE 426 AA; 47418 MW; DC8FA3DB4572BB40 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 426;

Best Local Similarity 71.4%; Pred. No. 9.7;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
Db 101 CKEDYR 107

RESULT 11
COLB_ARATH
ID COLB_ARATH STANDARD; PRT; 372 AA.
AC Q9SSES;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Zinc finger protein constans-like 11.
DE Zinc finger protein constans-like 11.
GN AT3G07650 OR MFP3.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salancoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,
RA Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA Delensy M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choisine N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barques M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan R., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argirion A., Flores M., Liguori R., Vitale D.,
RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Millischer J., Sellers P., Gill J.E., Feldblum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.
CC -!- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AC009176; AAF13083.1; -.
CC InterPro; IPR000315; Znf_Bbox.
CC InterPro; IPR002926; Znf_Constans.
CC Pfam; PF00643; zf-B_box; 2.
CC ProDom; PD007661; Znf_Constans; 1.
CC SMART; SM00336; BBOX; 2.

```

DR PROSITE; PS00119; ZF_BBOX; 2.
 KW Zinc-finger; Nuclear protein; Repeat; Multigene family.
 FT ZN_FING 5 47 B BOX-TYPE 1.
 FT ZN_FING 48 99 B BOX-TYPE 2 (ATPICAL).
 FT DOMAIN 77 83 POLY-ASN.
 FT DOMAIN 84 90 POLY-SER.
 SQ SEQUENCE 372 AA; 40754 MW; 188F18EB283D7479 CRC64;
 Query Match 81.4%; Score 35; DB 1; Length 372;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEEDFY 6
 ||:||||
 Db 181 CEDDFY 186
 RESULT 12
 BAL_MOUSE STANDARD; PRT; 599 AA.
 AC Q64285;
 01-NOV-1997 (Rel. 35, Created)
 01-NOV-1997 (Rel. 35, Last sequence update)
 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Bile-salt-activated lipase precursor (EC 3.1.1.13) (BAL)
 DE Bile-salt-activated lipase precursor (EC 3.1.1.13) (BAL)
 DE (Bile-salt-stimulated lipase) (BSSL) (Carboxyl ester lipase) (Sterol
 DE esterase) (Cholesterol esterase) (Pancreatic lysophospholipase).
 GN CEL OR Lip1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Mammary gland;
 RX MEDLINE=96096531; PubMed=8522186;
 RA Mackay K., Lawn R.M.;
 RT "Molecular cloning and characterization of the mouse carboxyl ester
 RT lipase gene and evidence for expression in the lactating mammary
 RT gland.";
 RL Genomics 29:115-122(1995).
 CC -1- FUNCTION: CATALYZES FAT AND VITAMIN ABSORPTION. ACTS IN CONCERT
 CC WITH PANCREATIC LIPASE AND COLIPASE FOR THE COMPLETE DIGESTION
 CC OF DIETARY TRIGLYCERIDES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: triacylglycerol + H(2)O = diacylglycerol + a
 CC fatty acid anion.
 CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U33169; AAA92088.1; -;
 CC EMBL; U37386; AAC52279.1; -;
 CC HSSP; P30122; 2BCE
 CC SWISS-2DPAGE; Q64285; MOUSE.
 CC MGD; MGI:88374; Cel.
 CC InterPro; IPR002018; CarboxylesteraseB.
 CC InterPro; IPR000379; Ser_estrs_site.

DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 KW Hydrolase; Serine esterase; Lipid degradation; Glycoprotein;
 KW Repeat; Signal.
 FT SIGNAL 1 20 BY SIMILARITY.
 FT CHAIN 21 599 BILE-SALT-ACTIVATED LIPASE.
 FT ACT_SITE 214 214 BY SIMILARITY.
 FT ACT_SITE 340 340 BY SIMILARITY.
 FT ACT_SITE 455 455 BY SIMILARITY.
 FT DISULFID 84 100 BY SIMILARITY.
 FT DISULFID 266 277 BY SIMILARITY.
 FT DOMAIN 559 588 4 X 11 AA TANDEM REPEATS, O-GLYCOSYLATED
 FT REPEAT 559 569 1.
 FT REPEAT 570 580 2.
 FT REPEAT 581 588 3.
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 599 AA; 65813 MW; 9E4428FDCA8602E CRC64;
 Query Match 79.1%; Score 34; DB 1; Length 599;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EEDFYR 7
 |||||
 Db 361 EEDFYR 366
 RESULT 13
 BAL_RAT STANDARD; PRT; 612 AA.
 AC P07882; P14722;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Bile-salt-activated lipase precursor (EC 3.1.1.13) (BAL)
 DE (Bile-salt-stimulated lipase) (BSSL) (Carboxyl ester lipase) (Sterol
 DE esterase) (Cholesterol esterase) (Pancreatic lysophospholipase).
 GN CEL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Pancreas;
 RX MEDLINE=90089378; PubMed=2688744;
 RA Kissel J.A., Fontaine R.N., Turk C.W., Brockman H.L., Hui D.Y.;
 RT "Molecular cloning and expression of cDNA for rat pancreatic
 RT cholesterol esterase.";
 RL Biochim. Biophys. Acta 1006:227-237(1989).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=87242339; PubMed=3593682;
 RA Han J.H., Stratowa C., Rutter W.J.;
 RT "Isolation of full-length putative rat lysophospholipase cDNA using
 RT improved methods for mRNA isolation and cDNA cloning.";
 RL Biochemistry 26:1617-1625(1987).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=91299758; PubMed=2069957;
 RA Fontaine R.N., Carter C.P., Hui D.Y.;
 RT "Structure of the rat pancreatic cholesterol esterase gene.";
 RL Biochemistry 30:7008-7014(1991).
 RN [4]
 RN ACTIVE SITE SER-214.
 RX MEDLINE=91009095; PubMed=2211595;
 RA Dipersio L.P., Fontaine R.N., Hui D.Y.;
 RT "Identification of the active site serine in pancreatic cholesterol
 RT esterase by chemical modification and site-specific mutagenesis.";
 RL J. Biol. Chem. 265:16801-16806(1990).

[5]
RN ACTIVE SITE HIS-455.
RX MEDLINE=91154187; PubMed=1999399;
RA Diersio L.P., Fontaine R.N., Hui D.Y.;
RT "Site-specific mutagenesis of an essential histidine residue in
pancreatic cholesterol esterase";
RL J. Biol. Chem. 266:4033-4036(1991).
CC -|- FUNCTION: CATALYZES FAT AND VITAMIN ABSORPTION. ACTS IN CONCERT
WITH PANCREATIC LIPASE AND COLIPASE FOR THE COMPLETE DIGESTION
OF DIETARY TRIGLYCERIDES.
CC -|- CATALYTIC ACTIVITY: Triacylglycerol + a
fatty acid anion.
CC -|- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
acid.
CC -|- ENZYME REGULATION: ACTIVATED BY BILE SALTS CONTAINING A 7-HYDROXYL
GROUP.
CC -|- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE PANCREAS AND THEN
TRANSPORTED TO THE INTESTINE.
CC -|- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X16054; CAA34189.1; -;
DR EMBL; M15893; AAA41540.1; -;
DR EMBL; M69157; AAB46376.1; -;
DR PIR; A34967; A34967.
DR PIR; A26603; A26603.
DR HSSP; P30122; 28CE.
DR InterPro; IPR002018; CarbesteraseB.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B.1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B.2; 1.
KW Hydrolase; Serine esterase; Lipid degradation; Glycoprotein;
KW Repeat; Signal.
KW SIGNAL 1 20
FT CHAIN 21 612 BILE-SALT-ACTIVATED LIPASE.
FT ACT_SITE 214 214
FT ACT_SITE 340 340
FT ACT_SITE 455 455
FT ACT_SITE 455 455
FT DISULFID 84 100
FT DISULFID 266 277
FT CARBOHYD 207 207
FT DOMAIN 556 599
FT REPEAT 556 566
FT REPEAT 567 577
FT REPEAT 578 588
FT REPEAT 589 599
FT MUTAGEN 440 440
FT MUTAGEN 455 455
FT MUTAGEN 455 455
FT CONFLICT 26 26
FT CONFLICT 154 154
FT CONFLICT 217 217
FT CONFLICT 219 219
FT CONFLICT 419 419
FT CONFLICT 513 513
FT CONFLICT 576 577
FT CONFLICT 608 609
FT CONFLICT 611 611
SQ SEQUENCE 612 AA; 67040 MW; 1569CE4EA71ED02A CRC64;
Query Match 79.1%; Score 34; DB 1; Length 612;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 BEDFYR 7

Db 361 BEDFYR 366
|||||
RESULT 14
NRDD_HAEIN
ID NRDD_HAEIN STANDARD; PRT; 707 AA.
AC P43752;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anaerobic ribonucleoside-triphosphate reductase (EC 1.17.4.2).
GN NRDD OR HI0075.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKennedy K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -|- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside triphosphate + oxidized
thioredoxin + H(2)O = ribonucleoside triphosphate + reduced
thioredoxin.
CC -|- SUBUNIT: Tetramer consisting of 2 alpha (NrdD) and 2 beta (NrdG)
subunits (By similarity).
CC -|- SIMILARITY: STRONG. TO E.COLI AND T4 NRDD.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U32693; AAC21751.1; -;
DR HSSP; P07071; 1B8B.
DR TIGR; HI0075; -;
DR InterPro; IPR005144; ATP_Actransf.
DR InterPro; IPR001150; Frm_Actransf.
DR Pfam; PF01228; Gly_radical; 1.
DR Pfam; PF03477; ATP-cone; 1.
DR PROSITE; PS00850; GLY_RADICAL; 1.
KW Oxidoreductase; Organic radical; Complete proteome.
FT MOD_RES 682 682 FREE RADICAL (BY SIMILARITY).
SQ SEQUENCE 707 AA; 80233 MW; A3795F7921A6781D CRC64;
Query Match 79.1%; Score 34; DB 1; Length 707;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 BEDFYR 7
|||||
Db 425 BEDFYR 430
RESULT 15
Y576_ARCFU
ID Y576_ARCFU STANDARD; PRT; 86 AA.
AC O29679;

```
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0576.
GN AF0576.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervilange A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RI "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001065; AAB90675.1; -
CC TIGR; AF0576; -
CC Hypothetical protein: Complete proteome.
KW SEQUENCE 86 AA; 10018 MW; 03AA3D19DAB8000C CRC64;
QY Query Match 76.7%; Score 33; DB 1; Length 86;
DR Best Local Similarity 71.4%; Pred. No. 7.1;
KW Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CEEDFYR 7
DB III: II
DB 67 CEEIYR 73
```

Search completed: February 12, 2003, 11:44:59
time : 7.53333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 11:42:45 ; Search time 52.7333 Seconds
(without alignments)
27.351 Million cell updates/sec

Title: US-09-660-302c-7
Perfect score: 43
Sequence: 1 CEEDFYR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|-------------|
| 1 | 38 | 88.4 | 153 | 16 | Q8R8G6 |
| 2 | 38 | 88.4 | 632 | 6 | Q95ML5 |
| 3 | 38 | 88.4 | 638 | 6 | Q9XSZ1 |
| 4 | 37 | 86.0 | 399 | 11 | Q9QV66 |
| 5 | 37 | 86.0 | 399 | 11 | Q8VCO7 |
| 6 | 36 | 83.7 | 76 | 5 | Q9U9B1 |
| 7 | 36 | 83.7 | 91 | 5 | Q16938 |
| 8 | 36 | 83.7 | 125 | 13 | Q98SF6 |
| 9 | 36 | 83.7 | 206 | 4 | Q14897 |
| 10 | 36 | 83.7 | 217 | 13 | Q98SF7 |
| 11 | 36 | 83.7 | 297 | 11 | Q9CSG0 |
| 12 | 36 | 83.7 | 378 | 13 | Q90881 |
| 13 | 36 | 83.7 | 400 | 13 | Q42108 |
| 14 | 36 | 83.7 | 726 | 10 | Q9M9Y2 |
| 15 | 35 | 81.4 | 202 | 16 | Q9PNX2 |
| 16 | 35 | 81.4 | 219 | 10 | Q8S056 |

| | | | | | |
|----|----|------|------|----|--------|
| 17 | 35 | 81.4 | 552 | 16 | Q8ZQE6 |
| 18 | 35 | 81.4 | 552 | 16 | Q8Z827 |
| 19 | 35 | 81.4 | 1376 | 5 | Q26637 |
| 20 | 35 | 81.4 | 1823 | 5 | Q26638 |
| 21 | 35 | 81.4 | 3198 | 5 | Q26639 |
| 22 | 34 | 79.1 | 262 | 5 | Q8SU06 |
| 23 | 34 | 79.1 | 277 | 10 | Q9SK11 |
| 24 | 34 | 79.1 | 331 | 16 | Q97ID4 |
| 25 | 34 | 79.1 | 420 | 16 | Q8RA45 |
| 26 | 34 | 79.1 | 563 | 16 | Q8YN67 |
| 27 | 34 | 79.1 | 592 | 11 | Q64571 |
| 28 | 34 | 79.1 | 599 | 11 | Q922R3 |
| 29 | 34 | 79.1 | 625 | 3 | Q94300 |
| 30 | 34 | 79.1 | 739 | 11 | Q63202 |
| 31 | 34 | 79.1 | 1007 | 12 | Q9QH62 |
| 32 | 33 | 76.7 | 98 | 5 | Q16939 |
| 33 | 33 | 76.7 | 156 | 5 | Q9W409 |
| 34 | 33 | 76.7 | 243 | 5 | Q95X85 |
| 35 | 33 | 76.7 | 272 | 16 | Q98PN8 |
| 36 | 33 | 76.7 | 441 | 12 | Q9VVJ2 |
| 37 | 33 | 76.7 | 496 | 5 | Q45368 |
| 38 | 32 | 74.4 | 60 | 4 | Q96JP7 |
| 39 | 32 | 74.4 | 200 | 4 | Q9UFD6 |
| 40 | 32 | 74.4 | 213 | 10 | Q9FP18 |
| 41 | 32 | 74.4 | 214 | 13 | Q8UVD2 |
| 42 | 32 | 74.4 | 226 | 17 | Q82WM1 |
| 43 | 32 | 74.4 | 249 | 16 | Q97TP4 |
| 44 | 32 | 74.4 | 267 | 5 | P90673 |
| 45 | 32 | 74.4 | 272 | 5 | Q8WTN7 |

ALIGNMENTS

RESULT 1

Q8R8G6 ID Q8R8G6 PRELIMINARY; PRT; 153 AA.
AC Q8R8G6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein TTE2033.
GN TTE2033.
OS Thermoanaerobacter tengcongensis
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MBAT / JCM11007;
RX MEDLINE=21992816; Pubmed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013153; AAM25210.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 153 AA; 18428 MW; 3B0F4ED430B1A0FA CRC64;

Query Match 88.4% Score 38; DB 16; Length 153;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFY 6
Db 21 CEEDFY 26

RESULT 2

Q95ML5 ID Q95ML5 PRELIMINARY; PRT; 632 AA.
AC Q95ML5;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Growth hormone receptor.
 OS Saimiri boliviensis (Bolivian squirrel monkey).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
 OX NCBI_TaxID=27679;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21265430; PubMed=11371582;
 RA Liu J.C., Nakova K.D., Atkins R.M., Gibson S., Li W.H.;
 RT "Episodic Evolution of Growth Hormone in Primates and Emergence of the
 RT Species Specificity of Human Growth Hormone Receptor.";
 RL Mol. Biol. Evol. 18:945-953(2001).
 DR EMBL; AF339061; AAK62288.1; -;
 DR InterPro; IPR002996; CRIA.
 DR InterPro; IPR003961; FN.III.
 DR InterPro; IPR003528; Hemtopoptn_L_F1.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 DR PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
 DR Receptor.
 SEQUENCE 632 AA; 70883 MW; 440817AF6277EDA3 CRC64;
 Query Match 88.4%; Score 38; DB 6; Length 632;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEEDFY 6
 Db 259 CEEDFY 264
 |||||
 RESULT 3
 O9XSZ1 PRELIMINARY; PRT; 638 AA.
 ID Q9XSZ1
 AC Q9XSZ1;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Growth hormone receptor.
 OS Papio anubis (olive baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Papio.
 OX NCBI_TaxID=9555;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99367319; PubMed=10425448;
 RZ Zogopoulos G., Nathanielsz P., Hendy G.N., Goodyer C.G.;
 RT "The baboon: a model for the study of primate growth hormone receptor
 RT gene expression during development.";
 RL J. Mol. Endocrinol. 23:67-75(1999).
 DR EMBL; AF150751; AAD39536.1; -;
 DR HSP; P10912; IAXI.
 DR InterPro; IPR002996; CRIA.
 DR InterPro; IPR003961; FN.III.
 DR InterPro; IPR003528; Hemtopoptn_L_F1.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 DR PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
 DR Receptor.
 SEQUENCE 638 AA; 71407 MW; 9E2500C8E303E420 CRC64;
 Query Match 88.4%; Score 38; DB 6; Length 638;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEEDFY 6
 Db 259 CEEDFY 264
 |||||

RESULT 4
 Q9QY66 PRELIMINARY; PRT; 399 AA.
 ID Q9QY66
 AC Q9QY66;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Cllorf5.
 DE Cllorf5.
 GN ORF6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20069078; PubMed=10602999;
 RA Lemmens I.H., Farnebo F., Piehl F., Merregaert J., Van de Ven W.J.M.,
 RA Larsson C., Kas K.;
 RT "Molecular characterization of human and murine cllorf5, a new member
 RT of the FAUNA gene cluster.";
 RL Mamm. Genome 11:78-80(2000).
 DR EMBL; AF119498; AAF23592.1; -;
 DR MGD; MGI:1352481; ORF6.
 DR SEQUENCE 399 AA; 43038 MW; 6BED852632747B54 CRC64;
 Query Match 86.0%; Score 37; DB 11; Length 399;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CEEDFY 7
 Db 41 CAEDFYR 47
 |||||
 RESULT 5
 Q8VCO7 PRELIMINARY; PRT; 399 AA.
 ID Q8VCO7
 AC Q8VCO7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Open reading frame 6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX Tissue=Liver;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC019440; AAH19440.1; -;
 DR SEQUENCE 399 AA; 42995 MW; B5CA88342B56932B CRC64;
 Query Match 86.0%; Score 37; DB 11; Length 399;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CEEDFY 7
 Db 41 CAEDFYR 47
 |||||
 RESULT 6
 Q9U9B1 PRELIMINARY; PRT; 76 AA.
 ID Q9U9B1
 AC Q9U9B1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Ascaris type serine protease inhibitor (Fragment).
 OS Ancylostoma ceylanicum.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=53326;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrison L.M., Cappello M.;
RT "The molecular cloning of an ascaris type serine protease inhibitor
from adult Ancylostoma ceylanicum hookworms.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF172653; AAD51336.1; -;
DR HSP; P56682; ICCV.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW protease.
KW NON_TER 1 1
FT SEQUENCE 76 AA; 8385 MW; D35FCE7C2088A53 CRC64;

Query Match 83.7%; Score 36; DB 5; Length 76;
Best Local Similarity 85.7%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CEEDFYR 7
||| |||
Db 44 CEEGYR 50

RESULT 7
Q16938
ID Q16938 PRELIMINARY; PRT; 91 AA.
AC Q16938;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Anti-coagulant protein C2 precursor (Fragment).
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98298519; PubMed=9634780;
RA Jespers L.S., Messens J.H., De Keyser A., Beckhout D.,
van den Brande I., Gansemans Y.G., Lauwereys M.J., Vlasuk G.P.,
Stanssens P.E.;
RT "Surface expression and ligand-based selection of cDNAs fused to
filamentous phage gene VI.";
RL Biotechnology 13:378-382(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312555; PubMed=8700900;
RA Stanssens P., Bergum P.W., Gansemans Y., Jaspers L., Laroche Y.,
Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,
Lasters I., Vlasuk G.P.;
RT "Anticoagulant repertoire of the hookworm Ancylostoma caninum.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).
DR EMBL; U30793; AAC47080.1; -;
DR HSP; P56682; ICCV.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW Signal.
KW NON_TER 1 1
FT SIGNAL <1 7
FT CHAIN 8 91
FT CHAIN 8 91
SQ SEQUENCE 91 AA; 10358 MW; ECB11CB4597C24DA CRC64;

Query Match 83.7%; Score 36; DB 5; Length 91;
Best Local Similarity 85.7%; Pred. No. 7.3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CEEDFYR 7
||| |||
Db 59 CEEGYR 65

RESULT 8
Q98SF6
ID Q98SF6 PRELIMINARY; PRT; 125 AA.
AC Q98SF6;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Lhx2 protein (Fragment).
GN LHX2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEAD;
RA Bachy I., Vernier P., Retaux S.;
RT "The LIM-homeodomain family in the developing xenopus brain:
conservation and divergences with the mouse related to the evolution
of the forebrain.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
IONS.
DR EMBL; AJ311712; CAC35215.1; -;
DR HSP; P32965; ICTL.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 2.
DR ProDom; PD000094; LIM; 1.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00478; LIM_DOMAIN_1; 1.
DR PROSITE; PS50023; LIM_DOMAIN_2; 2.
KW LIM domain; Metal-binding; Zinc.
FT NON_TER 1 1
FT NON_TER 125 125
SQ SEQUENCE 125 AA; 14283 MW; 375E42A29104D364 CRC64;

Query Match 83.7%; Score 36; DB 13; Length 125;
Best Local Similarity 71.4%; Pred. No. 10;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CEEDFYR 7
||| |||
Db 28 CKEDYR 34

RESULT 9
Q14897
ID Q14897 PRELIMINARY; PRT; 206 AA.
AC Q14897;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE L-myc protein (Similar to lung carcinoma myc related oncogene 1).
GN L-MYC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=88094386; PubMed=2827002;
RA Kaye F., Battey J., Nau M., Brooks B., Seifter E., De Greve J.,
Birrer M., Sausville E., Minna J.;
RT "Structure and expression of the human L-myc gene reveal a complex
pattern of alternative mRNA processing.";
RL Mol. Cell. Biol. 8:186-195(1988).

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M19720; AAAS9878.1; -;
 DR EMBL; BC011864; AAH11864.1; -;
 DR TRANSFAC; T02386; -;
 DR InterPro; IPR002418; TF_Myc.
 DR Pfam; PF01056; MYC_N term; 1.
 SQ SEQUENCE 206 AA; 21766 MW; 585C9CD6C9A8EC71 CRC64;

Query Match 83.7%; Score 36; DB 4; Length 206;
 Best Local Similarity 85.7%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
 I : I I I I I
 DB 15 CGEDFYR 21

RESULT 10

Q98SF7 PRELIMINARY; PRT; 217 AA.
 Q98SF7; (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Lhx9 protein (Fragment).
 GN LHX9.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEAD;
 RA Bachy I., Vernier P., Retaux S.;
 RT "The LIM-homeodomain family in the developing xenopus brain:
 conservation and divergences with the mouse related to the evolution
 of the forebrain."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2 ZINC
 IONS.
 DR EMBL; AJ311171; CAC35214.1; -;
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00412; LIM; 2.
 DR ProDom; PD000010; Homeobox; 1.
 DR ProDom; PD000094; LIM; 2.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00132; LIM; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 1.
 DR PROSITE; PS50023; LIM_DOMAIN_2; 1.
 KW LIM domain; Metal-binding; Zinc.
 FT NON_TER 1
 FT NON_TER 217
 SQ SEQUENCE 217 AA; 24716 MW; 6CBE88B595ECB851 CRC64;

Query Match 83.7%; Score 36; DB 13; Length 217;
 Best Local Similarity 71.4%; Pred. No. 18;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
 I : I I I I I
 DB 27 CKEDYR 33

RESULT 11

Q9CSGO PRELIMINARY; PRT; 297 AA.

ID Q9CSGO PRELIMINARY; PRT; 297 AA.
 AC Q9CSGO;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE LIM homeobox protein 9 (Fragment).
 GN LHX9.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 IONS.
 DR EMBL; AK012930; BAB28555.1; -;
 DR HSP; P32965; ICTL.
 DR MGD; MGI:1316721; Lhx9.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00412; LIM; 2.
 DR ProDom; PD000010; Homeobox; 1.
 DR ProDom; PD000094; LIM; 2.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00132; LIM; 2.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 1.
 DR PROSITE; PS50023; LIM_DOMAIN_2; 2.
 KW LIM domain; Metal-binding; Zinc.
 FT NON_TER 1
 FT NON_TER 297
 SQ SEQUENCE 297 AA; 33502 MW; 3840FE3B819E8053 CRC64;

Query Match 83.7%; Score 36; DB 11; Length 297;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEEDFYR 7
 I : I I I I I
 DB 87 CKEDYR 93

RESULT 12

Q90881 PRELIMINARY; PRT; 378 AA.
 ID Q90881
 AC Q90881;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homeobox protein.
 GN LH-2.

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Tremml G., Jessell T.M.;
RT "Differentiation of dorsal commissural neurons defined by expression
RT of the LIM homeobox gene Lh-2: Suppression by notochord grafts and
RT maintained after notochord removal.";
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL; L35566; AAA50258.1; -.
DR HSSP; P32965; ICTL.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_repressr.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF04112; LIM; 2.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000094; LIM; 2.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS50023; LIM_DOMAIN_2; 2.
KW DNA-binding; Homeobox; LIM domain; Metal-binding; Nuclear protein;
KW Zinc.
SQ SEQUENCE 378 AA; 42007 MW; 34220850FCB82FFC CRC64;
Query Match 83.7%; Score 36; DB 13; Length 378;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEDFYR 7
Db 101 CKEDYR 107
I:|:|:|
|:|:|:|
RESULT 13
O42108
ID O42108 PRELIMINARY; PRT; 400 AA.
AC O42108;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Lh-2A.
GN Lh-2A.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIMB BUD;
RX MEDLINE=97446014; PubMed=9299541;
RA Nohno T., Kawakami Y., Wada N., Ishikawa T., Ohuchi H., Noji S.;
RT "Differential expression of the two closely related LIM-class homeobox
RT genes Lh-2A and Lh-2B during limb development.";
RL Biochem. Biophys. Res. Commun. 238:506-511(1997).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL; AB005882; BAA21846.1; -.
DR HSSP; P06601; IFJL.
DR InterPro; IPR001356; Homeobox.

DR InterPro; IPR001781; LIM.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF00412; LIM; 2.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000094; LIM; 2.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS50023; LIM_DOMAIN_2; 2.
KW DNA-binding; Homeobox; LIM domain; Metal-binding; Nuclear protein;
KW Zinc.
SQ SEQUENCE 400 AA; 44339 MW; F687E764F233CD4C CRC64;
Query Match 83.7%; Score 36; DB 13; Length 400;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEDFYR 7
Db 103 CKEDYR 109
I:|:|:|
|:|:|:|
RESULT 14
Q9M9Y2
ID Q9M9Y2 PRELIMINARY; PRT; 726 AA.
AC Q9M9Y2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE F4H5.17 protein.
GN F4H5.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luross S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011001; AAF63144.1; -.
DR InterPro; IPR004332; MuDR.
DR InterPro; IPR004862; MurA.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF03108; MuDR; 1.
DR Pfam; PF03163; MURA; 1.
DR Pfam; PF00098; Zf-CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
SQ SEQUENCE 726 AA; 82359 MW; 6B9A6133C3BFF5B2 CRC64;
Query Match 83.7%; Score 36; DB 10; Length 726;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CEDFYR 7
Db 633 CKEDYR 639
I:|:|:|
|:|:|:|
RESULT 15
Q9PNX2
ID Q9PNX2 PRELIMINARY; PRT; 202 AA.
AC Q9PNX2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Hypothetical protein Cj0963.

Thu Feb 13 09:30:37 2003

GN CJ0963.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; ALI39076; CAB3220.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 202 AA; 23926 MW; BFB9C487F8642706 CRC64;

Query Match 81.4%; Score 35; DB 16; Length 202;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEDFY 6
I:||||
Db 99 CEDFY 104

Search completed: February 12, 2003, 11:47:00
Job time : 54.7333 secs